Modular Bayesian uncertainty assessment for structural health monitoring

by

André Henriques Jesus

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

School of Engineering

June 2018

Contents

Acknowledgments

Thesis funding

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) reference number EP/N509796.

Close collaborators/support

Firstly, I would like to acknowledge the help and support from my supervisors Dr Irwanda Laory, Dr Peter Brommer and Professor Toby Mottram. Ranging from the development of collaborations, the most theoretical aspects, or to the most pertinent details of the current thesis, their influence and kindness was astonishing. I also want to express my gratitude to Dr Stana Zivanovic and Professor Keith Worden, for the rigorous review of the present thesis, and the numerous pertinent comments which have enriched greatly its contents.

To my colleagues and friends Dr Mohammad Reza Salami, Arya Pamuncak and Yanjie Zhu, who have kindly shared countless hours of their experience in structural health monitoring with me, at both practical and theoretical level. All of them went beyond any call of duty (specially during laboratory experiments).

I benefited immensely from my experience with several personnel from the School of Engineering staff, students among other exceptional individuals. Hence, I want to thank Lee Davis who has been an extraordinary mentor at all levels since my come at the University of Warwick. My final load tests of a linear spring have also been performed successfully thanks to his immense patience and experience. I want

to acknowledge the outstanding contribution of my friends Neil Gillespie, Arnett Taylor and Colin Banks to this thesis. Their vast technical experience was essential during the tests of the aluminium bridge, the spring load tests, the IMC bridge instrumentation and endless other occasions. Finally, students Hannah Davies, Zong Tan, Hassib Abdulrazaq, and staff members Philip Moore, Adrian Seymour, Roger Thorpe, Jonathan Meadows & William Higgins rigour and exceptional professionalism have also been key to the success of this work.

Secondly, my sincere gratitude is directed to Professor James Brownjohn, Dr Ki-Young Koo and Dr Robert Westgate from the University of Exeter. I felt immensely privileged for having been able to collaborate with this group. Namely, due to a large amount of monitored data from the Tamar bridge which has been supplied, along with its complex finite element model, and the key discussions for validation of results.

Thirdly, to Zuzana Dimitrovova, Manuel Gonçalves da Silva, Richard & Diana Organ, without whom this thesis could not even have started, I wish to express my deepest thanks. I had the absolute honour of meeting my friend and tennis mentor Dr Ninoslav Pesic (Nino), with whom I had so many exciting matches and conversations. Also to my great judo fellows who have thrown and been thrown by me so many times, Lukasz Maliczenko, Harvey Lewis, Rouni, Ian Flemming, Toby, John and Matt. I want to thank to my friend Petros Kontos, for his awesome personality, friendly nature and sharing of our quad's fetish and DIY spirit. The sole names Michael Wingens and Maria Vittoria (Mavi) revive so many memorable moments during my stay at Claycroft 2. Finally, Ren, Shawn Jing, Ercan Dietzfelbinger, Olumese, Li Ruowei, Mu, Liam Mosley and everyone from International House, have widened immensely my world view. I still vividly recall all the extraordinary moments lived during a single academic year.

Finally, I hereby express my gratitude to my family. To my first cousin once removed and second cousin, Anna and Mina, who so warmly have welcomed me during my stay in the land of Oz. My brother and sister, Ricardo and Sofia, who always challenge and motivate me to be a better person. To my parents, Ana Estela

and Sergio Jesus for their ever-constant love, always instilling their perfectionism and opening my world to new experiences. To my dear Carla, for her patience, belief and love throughout this entire project. Her influence was profound, and I cannot convey a sufficient level of admiration and appreciation for her support.

Software & Hardware

- Most of the models in this thesis were developed with ANSYS Inc. A student license of MATLAB was also used to develop the main methodology. The author gratefully acknowledges the computing resources provided by the Cluster of Workstations operated by the Centre for Scientific Computing of the University of Warwick. Finally all editing, writing, planning and coding was centred around the GNU Emacs text editor.

Declarations

This thesis is based upon material from the following publications

- Jesus, A., Brommer, P., Zhu, Y. AND Laory, I. Comprehensive Bayesian Structural Identification Using Temperature Variation. *Engineering Structures* 141 (June 2017), 75 –82. ISSN: 0141-0296. DOI: 10.1016/j.engstruct.2017.01.060.

- Jesus, A., Brommer, P., Westgate, R., Koo, K.-Y., Brownjohn, J. M. W. AND Laory, I. Bayesian Structural Identification of a Long Suspension Bridge Considering Temperature and Traffic Load Effects. *in review Journal of Structural Health Monitoring* (20th Mar. 2018).

- Jesus, A., Zhu, Y. AND Laory, I. Comprehensive Bayesian Structural Identification Using Temperature Expansion of a Scale Aluminium Bridge. In: *8th European Workshop on Structural Health Monitoring, EWSHM 2016*. Vol. 4. 2016, 3080–3088.

- Jesus, A., Salami, M., Westgate, R., Koo, K.-Y., Brownjohn, J. AND Laory, I. Bayesian Structural Identification of a Suspension Bridge Using Temperature and Traffic Loading. In: *Proceedings of the 8th International Conference on Structural Health Monitoring of Intelligent Infrastructure (SHMII), Brisbane*. 2017.

- Jesus, A., Zhu, Y., Koo, K.-Y., Brownjohn, J. AND Laory, I. Identifiability Based Sensor Configuration Analysis for Bayesian Structural Identification. In:

Proceedings of the 8th International Conference on Structural Health Monitoring of Intelligent Infrastructure (SHMII), Brisbane. 2017.

Abstract

Civil infrastructure are critical elements to a society welfare and economic thriving. Understanding their behaviour and monitoring their serviceability are relevant challenges of Structural Health Monitoring (SHM). Despite the impressive improvement of miniaturisation, standardisation and diversity of monitoring systems, the ability to interpret data has registered a much slower progression across years. The underlying causes for such disparity are the overall complexity of the proposed challenge, and the inherent errors and lack of information associated with it. Overall, it is necessary to appropriately quantify the uncertainties which undermine the SHM concept.

This thesis proposes an enhanced modular Bayesian framework (MBA) for structural identification (st-id) and measurement system design (MSD). The framework is hybrid, in the sense that it uses a physics-based model, and Gaussian processes (mrGp) which are trained against data, for uncertainty quantification. The mrGp act as emulators of the model response surface and its model discrepancy, also quantifying observation error, parametric and interpolation uncertainty. Finally, this framework has been enhanced with the Metropolis–Hastings for multiple parameters st-id. In contrast to other probabilistic frameworks, the MBA allows to estimate structural parameters (which reflect a performance of interest) consistently with their physical interpretation, while highlighting patterns of a model's discrepancy. The MBA performance can be substantially improved by considering multiple responses which are sensitive to the structural parameters.

An extension of the MBA for MSD has been validated on a reduced-scale aluminium bridge subject to thermal expansion (supported at one end with springs and instrumented with strain gauges and thermocouples). A finite element (FE) model of the structure was used to obtain a semi-optimal sensor configuration for st-id. Results indicate that 1) measuring responses which are sensitive to the structural parameters and are more directly related to model discrepancy, provide the best results for st-id; 2) prior knowledge of the model discrepancy is essential to capture the latter type of responses. Subsequently, an extension of the MBA for st-id was also applied for identification of the springs stiffness, and results indicate relative errors five times less than other state of the art Bayesian/deterministic methodologies.

Finally, a first application to field data was performed, to calibrate a detailed FE model of the Tamar suspension bridge using long-term monitored data. Mea-

surements of temperature, traffic, mid-span displacement and natural frequencies of the bridge, were used to identify the bridge's main/stay cables initial strain and friction of its bearings. Validation of results suggests that the identified parameters agree more closely with the true structural behaviour of the bridge, with an error that is several orders of magnitude smaller than other probabilistic st-id approaches. Additionally, the MBA allowed to predicted model discrepancy functions to assess the predictive ability of the Tamar bridge FE model. It was found, that the model predicts more accurately the bridge mid-span displacements than its natural frequencies, and that the adopted traffic model is less able to simulate the bridge behaviour during periods of traffic jams.

Future developments of the MBA framework include its extension and application for damage detection and MSD with multiple parameter identification.

List of Tables

List of Figures

Nomenclature

Roman symbols

$\boldsymbol{D}^m$ simulation dataset, page 74

$\boldsymbol{C}, \boldsymbol{C}_m, \boldsymbol{C}_e, \boldsymbol{C}_\delta, \boldsymbol{C}_Y$ Lower triangular Cholesky decomposition of correlation or covariance matrix generic/model/experiment/discrepancy function/global model, page 64

$\boldsymbol{D}$ combined/generic dataset, page 80

$\boldsymbol{D}^e$ experimental dataset, page 77

$\boldsymbol{G}, \boldsymbol{G}_m, \boldsymbol{G}_\delta, \boldsymbol{G}_e\boldsymbol{G}_Y$ upper triangular matrix resulting from QR factorisation generic/ model/discrepancy function/observations/global model, page 176

$\boldsymbol{h}_Y, \boldsymbol{H}_Y$ hierarchy of linear regression functions for combined data vector/matrix, page 82

$\boldsymbol{m}, \boldsymbol{m}_e, \boldsymbol{m}_Y$ mean vector generic/of observations/combined data, page 50

$\boldsymbol{Q}, \boldsymbol{Q}_m, \boldsymbol{Q}_\delta$ orthogonal matrix of QR factorisation generic/model/discrepancy function, page 176

$\boldsymbol{V}, \boldsymbol{V}_m, \boldsymbol{V}_e, \boldsymbol{V}_Y$ covariance matrix of model/experiment/combined data, page 53

$\boldsymbol{W}_\delta$ Gram matrix inverse of discrepancy function, page 80

$\boldsymbol{W}_m$ Gram matrix inverse of model, page 58

$\boldsymbol{W}_Y$ Gram matrix inverse of combined data, page 83

$\boldsymbol{Y}^\delta$ discrepancy function dataset, page 178

$\hat{\boldsymbol{Y}}^m$ integrated model dataset, page 79

$\mathcal{R}, \mathcal{R}_m, \mathcal{R}_\delta$ relational correlation vector/matrix for model/discrepancy function, page 59

$\mathcal{V}$ relational covariance vector/matrix, page 63

$\widetilde{\boldsymbol{H}}, \widetilde{\boldsymbol{H}}_m, \widetilde{\boldsymbol{H}}_\delta$ Cholesky factorised matrix of linear regression functions generic/ model/discrepancy function, page 64

$\widetilde{\boldsymbol{Y}}$, $\widetilde{\boldsymbol{Y}}^m$, $\widetilde{\boldsymbol{Y}}^\delta$ Choleksy transformed matrix of process output generic/model/discrepancy function, page 64

d number of design variables, page 67

$h, \boldsymbol{h}, \boldsymbol{H}$ hierarchy of linear regression functions, page 50

$h_\delta, \boldsymbol{h}_\delta, \boldsymbol{H}_\delta$ hierarchy of linear regression functions of discrepancy function, page 78

$h_m, \boldsymbol{h}_m, \boldsymbol{H}_m$ hierarchy of linear regression functions of model, page 75

N number of simulated data points, page 74

n number of observations, page 67

q number of responses, page 67

r number of structural parameters, page 68

$R, \boldsymbol{R}$ correlation function/matrix, page 51

$R_m, \boldsymbol{R}_m$ correlation function/matrix of model, page 75

S number of samples drawn with the Metropolis–Hastings algorithm, page 89

s number of predicted data points, page 75

$x, \boldsymbol{x}, \boldsymbol{X}$ design variables (generic), scalar/vector/matrix, page 68

$x^e, \boldsymbol{x}^e, \boldsymbol{X}^e$ measured design variables of observations, scalar/vector/matrix, page 68

$x^m, \boldsymbol{x}^m, \boldsymbol{X}^m$ design variables input of physics based model, scalar/vector/matrix, page 74

x_*, θ_* unobserved new input data points for predictions, page 58

$y, \boldsymbol{y}, \boldsymbol{Y}$ process output, scalar, vector matrix, page 48

$y^e, \boldsymbol{y}^e, \boldsymbol{Y}^e$ measured observations, scalar, vector matrix, page 68

$y^m, \boldsymbol{y}^m, \boldsymbol{Y}^m$ simulations from physics-based model scalar/vector/matrix, page 68

Greek symbols and maths

$\boldsymbol{\theta}^*$ true parameters value, page 69

$\boldsymbol{\varepsilon}$ observation error, page 68

$\boldsymbol{\beta}$ regression coefficients, page 50

$\boldsymbol{\beta}_\delta$ regression coefficients of discrepancy function, page 79

$\boldsymbol{\beta}_m$ regression coefficients of model, page 75

$\boldsymbol{\delta}$ discrepancy function, page 68

$\boldsymbol{\Lambda}$ variance of observation error, page 68

$\boldsymbol{\omega}, \boldsymbol{\omega}_m, \boldsymbol{\omega}_x, \boldsymbol{\omega}_\theta, \boldsymbol{\omega}_\delta$ roughness parameters of model/design variables/structural parameters/discrepancy function, page 52

$\boldsymbol{\phi}, \boldsymbol{\phi}^m, \boldsymbol{\phi}^\delta$ hyperparameters of mrGp model/discrepancy function, page 49

$\boldsymbol{\Sigma}^2, \boldsymbol{\Sigma}_m^2, \boldsymbol{\Sigma}_\delta^2$ response variance matrix of model/discrepancy function, page 53

$\boldsymbol{\theta}$ structural parameters generic value, page 68

σ^2 scalar response variance, page 51

ξ true unobservable process, page 67

$p(\boldsymbol{D}|\boldsymbol{\theta})$ likelihood function, page 67

$p(\boldsymbol{\theta})$ prior distribution, page 67

$p(\boldsymbol{\theta}|\boldsymbol{D})$ posterior distribution, page 67

Abbreviations

FE finite element, page 3

MBA Modular Bayesian approach, page 38

MCMC Markov chain Monte Carlo, page 22

MH Metropolis–Hastings algorithm, page 89

MLE Maximum likelihood estimation, page 49

mrGp, mrGp_m, mrGp_δ multiple response Gaussian process (of model/discrepancy function), page 48

MSD measurement system design, page 5

PDF probability density function, page 12

SHM Structural Health Monitoring, page 1

SNR signal-to-noise ratio, page 61

st-id structural identification, page 4

Chapter 1

Introduction

1.1 Thesis outline and problem statement

The inherent multidisciplinary nature required to develop a Structural Health Monitoring (SHM) process, combined with the economic, life-safety advantages provided by this technology, and its broad applications, qualifies it as a "Grand Challenge" problem for engineering in the 21st century.

From an economic perspective an SHM system represents 0.5% of a typical full-scale structure total cost, and its benefits comparatively to traditional inspections grossly justify its deployment [1]. As a consequence, at the beggining of 2016 the global SHM market is evaluated at £1243.6 millions [2] and is expected to register a compound annual growth rate (CAGR) of 13.8% over a ten year forecast period 2016–2026. These values are justified by a considerable progress in the standardisation and miniaturisation of SHM systems.

Regarding life-safety aspects, SHM replaces the conventional visual inspection-based decision-support with automated performance-based maintenance, drastically minimising the human involvement, and consequently reducing labour, downtime and human errors, and thus improving safety and reliability. The improvement of safety is a strong motivation, specially due to the problem of ageing structures. For example, bridge inspection reports worldwide have shown concerning ratios (up to 50%) of bridges which have deficiencies and structural anomalies not acceptable by present day standards.

Analogous to the human nervous system, an SHM process has two main components: a sensory system, represented by physical nodes located at key locations of a structure; and a central processing unit, which collects and interprets all available data (Fig. 1.1). In addition to typically measured structural responses e.g.

displacement, acceleration, strain, the sensory system is often designed to also record environmental/operational effects that a structure is subject to e.g. wind, temperature, loads. All of this data has to be analysed by the central processing unit, which has to extract relevant features and patterns, subsequently determining performance indices and other valuable decision-making information.

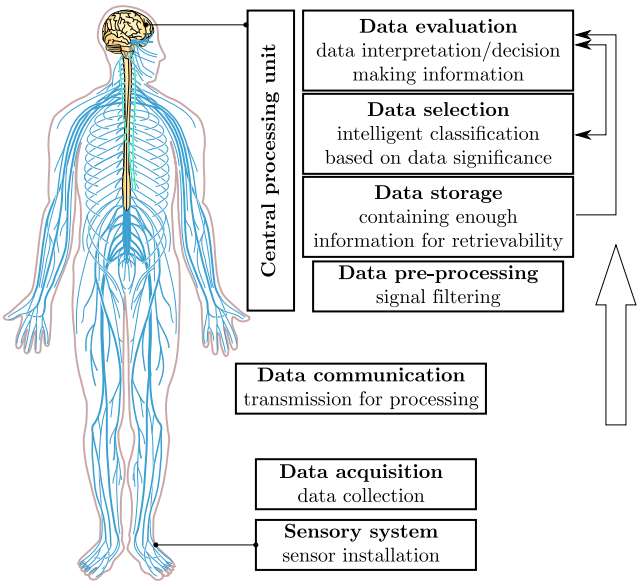


Figure 1.1: Flowchart of an effective SHM system (nervous system analogy).

The reliability and diversity of sensor technology, e.g. usage of fiber optic, micro electro mechanical sensors (MEMS), wireless data transmission and the ever presence of the internet, has increased considerably in recent decades, contributing to the widespread application of SHM in civil engineering. Despite such improvements, SHM systems are not yet as ubiquitous as expected.

One of the main hurdles against a more generalised use of SHM in the civil engineering industry is the gap between monitored data and actionable information required by stake-holders and decision-makers [3]. Such information includes performance indexes and/or reliability assessments which motivate additional visual inspections and repair/maintenance interventions. In essence, there is a large disparity between the amount of available data and the current level of interpretative ability. A common statement is that 'the world is data rich but information poor'.

Other problems are the inherent difficulties associated with data interpretation; which include handling of heterogeneous data, blending between monitored data/simulations, selection of appropriate data filtering or pattern recognition modules. Thus, validation of methodologies for the purpose of data interpretation is

more challenging, and standardisation protocols are less obviously established.

Finally, it should be noted that the SHM concept has a wide range of applications, such as aircraft, defence systems, rotorcraft or civil infrastructure (each with an associated degree of success). These systems justify deployment of long-term monitoring systems to assess their service life [4], due to their relevance for social welfare and economic growth. Civil infrastructure systems are particularly challenging, because of their large scale and uniqueness. For example, bridges are one-of-a-kind prototypes, and unlike disciplines of aeronautical and mechanical engineering, databases with variability of their structural components are nonexistent [5, 6].

1.2 Data interpretation mechanisms

The previous section introduced some core challenges of data interpretation in civil structures. Fundamentally, the problem of data interpretation consists in fitting a statistical model to the measured system response data, and using its logical workflow to highlight correlations, recognise patterns and establish dependencies between different quantities. Loosely speaking, a suitable 'structure' is assigned to the data, in order to understand its behaviour. Depending on the type of models that are used, three mechanisms or approaches can emerge:

- *physics-based models*, whose internal relations and parameters are based on the physical laws that govern the structural system, e.g. the finite element (FE) method or bond graph models. This type of interpretation is also known as model-based approach, mainly because it is more intuitively associated with the concept of a model;

- *data-based models*, which are based on machine learning algorithms not associated with the physical description of the system. Parameters of these models also do not have a physical interpretation. Examples of data-based models include cluster analyses, neural networks, or principle component analysis based methods. This type of interpretation is also known as model-free approach;

- *hybrid models* (also called integrated models) [7], are those that combine the two aforementioned approaches. Several authors defend that integrated approaches, which gather the joint effects of data and physics based models, provide the best results [8].

Data-based models explore the heterogeneity of data sets shaping different regions according to a well established criteria. Subsequently, the mapped regions

have to be associated with features of the structural system, in order to understand the underlying physical behaviour. On the other hand, physics-based models are constrained by the laws which govern the structural system, and how appropriately these conform to the monitored data. However, as long as proper calibration and validation has been carried out, the ability of these models to classify, and extrapolate beyond observed data is a major asset for data interpretation. Moreover, their development and use agrees with engineering knowledge, and is therefore more understandable. In the remainder of this section, three fields of application based on data interpretation, namely damage detection, structural identification and measurement system design, shall be highlighted.

1.2.1 Damage detection

Damage detection is the most relevant area of research in SHM, standing in contrast with classical visual inspection-based decision-support [7]. Development of early stage damage detection systems is a major challenge for acceptance of SHM by civil infrastructure owners. In the current context damage is defined as changes introduced into a structural system which adversely affect its current or future performance. An automated decision-making process for damage detection can be posed as a statistical pattern recognition paradigm. As posed by Rytter [9] and other authors [5, 10] the four levels of a damage detection hierarchy are: (1) detection, (2) location, (3) extent, (4) and prognosis of existent damage. It should be noted that due to the inability of data-based models to extrapolate data trends and generate predictions, only physics-based models are able to reach the prognostic level of the hierarchy.

1.2.2 Structural identification

Structural identification (st-id) is the process of creating, calibrating and updating a physics-based model of a structure based on monitored data. The calibrated model will subsequently be used for assessment of a structure's health, performance or decision making [11, 12]. There are deterministic and probabilistic methodologies used for st-id, and the latter are usually more realistic and well suited for continuous monitoring. Some examples of sources of uncertainty that should be considered in a probabilistic st-id process include, mechanical and geometric heterogeneity, non-stationary boundary properties, changing behaviour of moving systems, elevated intrinsic forces e.g., dead weight, uniqueness of the structural system, non-stationarity due to environmental and operational conditions. In short, st-id aims to estimate pa-

rameters and establish the correlation between a physics-based model and an actual structure.

Modal identification

A sub-class of the st-id area which is given particular emphasis is called modal identification. Since mode shapes and natural frequencies are seen as the most reliable indicator of damage, it is not surprising that modal identification assumes a significant relevance in the SHM community. This is so, because historically natural frequencies of rotary systems, such as helicopters blades, have proven to be features very susceptible to damage [5, 13–16]. As a consequence, an array of vibration-based methods have been developed in order to emulate the same approach to civil infrastructure. Unfortunately, unless under specific conditions, complex structures have proven to be not sensitive to this approach. This topic will be detailed more thoroughly in the literature review, chapter 2.

1.2.3 Measurement system design

Finally, measurement system design (MSD) aims to determine a sensor configuration of a monitoring system which allows to optimally interpret data, be it for the purpose of st-id or damage detection. Normally only simulated data is used to establish the configuration of the sensory system (since this is a task preliminary to the installation of the actual monitoring system). In addition to the key locations where sensors ought to be placed, what is actually measured is also relevant e.g. for the purpose of data fusion. Results typically indicate the *best* sensor configuration out of a set of available key locations, since it is unfeasible to consider *every* possible configuration. Generally MSD is a nonlinear and ill conditioned problem. It is nonlinear because the sensor configuration depends upon a performance criterion which in turn depends on the sensor configuration. It is ill-conditioned because the infinite number of possible sensor configurations brings an inherent ill conditioning of observations.

It should also be noted, that most of current SHM practice for MSD relies most times on specialist heuristics, rather than rigorous theory, often introducing biases and/or redundant information.

1.3 Scientific question and objectives

As detailed in the previous sections, the problem of data-interpretation is the core bottleneck of the SHM concept. There are numerous fields of SHM which address specific objectives of data interpretation, each with its own inherent difficulties.

The common conditioner to any data-interpretation mechanism or SHM field is the ubiquitous occurrence of *uncertainty*, in the form of errors and/or lack of information. Such factors severely compromise the reliability of interpreted actionable information. For these reasons, analysts often struggle to establish a decision-making basis for proper allocation of resources, e.g. for visual inspections and repair/ maintenance interventions.

As a consequence, the main problem which is to be addressed by the current thesis is the quantification of uncertainties which undermine an SHM process. Specifically, the current thesis will be focused on *a hybrid type of data interpretation which resorts to a combined physics-based and data model*, to assess unfavourable uncertainties. A robust probabilistic framework which accounts for a wide number of uncertainties will be developed and the advantages that it brings to the SHM state of the art will be highlighted [17].

The following scientific questions are considered:

- How prevalent are uncertainties in SHM and how do they affect data interpretation? In this context uncertainties are understood as any quantifiable error that negatively affects an SHM process.
- Having accepted the existence of uncertainties, how do their propagation and interaction affect the applicability of the SHM concept?
- Is it possible to mitigate, or at least quantify these errors in a rigorous way, so that outcomes of SHM might be improved?

The novel methodology proposed in this thesis is based on Bayesian probabilities. It is able to overcome normal drawbacks of state of the art methodologies due to its hybrid nature. Although the applicability of the methodology is extensive, the current work is restricted to the improvements that are brought to st-id. Furthermore, a MSD module which maximises the efficiency of the improved st-id approach will be developed.

Henceforth, the objectives of the work in this thesis are to:

- Develop and apply a state-of-the-art uncertainty quantification framework for SHM applications. The framework should be comprehensive towards the uncertainties it considers, and improve the application of SHM comparatively to previous works in scientific literature.
- Extend the developed framework for the purpose of MSD and st-id, illustrating the advantages that they bring by use of an enhanced uncertainty quantification framework.

- Validate the framework with simulated, reduced scale and full-scale case studies to ensure that the methodologies work as expected. A multitude of scenarios with different degrees of difficulty should be tested, in order to highlight the advantages, disadvantages and limitations of the proposed approach.

Finally, it is important to frame these objectives against model updating and the applicability of SHM. Model updating means the refinement of a model to improve the correlation between measurements and simulations. It is a natural follow-up task for the proposed uncertainty quantification, since the latter acts as a model performance classifier and updating operations guide. Relatively to the applicability of the present work during SHM, there are two relevant stages. A preliminary stage, when a monitoring system is to be designed and installed, which falls under the umbrella of MSD; and a middle stage, when monitored data has been collected and requires evaluation to assess the adequate performance of the structural system, which is the subject of st-id.

1.4 Organisation of text

The thesis is broken into seven distinct chapters. The layout is described below and a diagram of its contents is shown in Fig. 1.2.

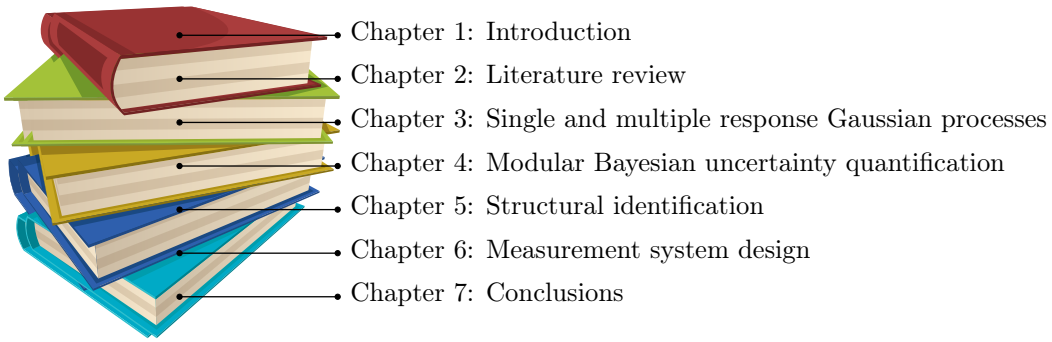


Figure 1.2: Thesis organisation (bullet point list).

Chapter 2 starts with a nomenclature of sources of uncertainty and core concepts of uncertainty quantification in order to establish an appropriate background. Subsequently this chapter presents a comprehensive literature review of the progress registered in uncertainty quantification of SHM, with particular emphasis on Bayesian methodologies. This chapter justifies the relevance of uncertainty quantification in SHM, and contextualises the objectives and goals of the work developed in this thesis.

Chapter 3 highlights a theoretical formulation of single and multiple response Gaussian processes, which is essential in order to familiarise the reader with the core methodology of the thesis. Where appropriate, differences between the single and multiple response case will be discussed. Data sets with and without noisy observations will be considered. Finally, a novel enhancement of the multiple response Gaussian process, authored by the thesis proponent, will be detailed in the last section of the chapter;

In **Chapter 4** the core methodology of this thesis is rigorously described. An explanatory data model is presented and interpreted on the basis of Bayes' theorem. The proposed framework is then expanded, with an implementation based on the Metropolis–Hastings algorithm authored by the thesis proponent. The last section of Chapter 4 showcases the application of the methodology to a group of simulated examples with different degrees of complexity. Some complementary details related with the formulation and computational aspects of the methodology are also described in appendix B and C;

Chapter 5 details the application of the methodology presented in the previous chapter for the purpose of structural identification. A case-study of an aluminium bridge and of a long suspension bridge are used to illustrate the application of the identification process to a reduced and full-scale infrastructure. Two FE models of these structures are presented, as well as their monitoring systems. Finally the results of identified model parameters, inadequacy between the models and actual structures, experimental errors and probabilistic predicted responses are also shown;

Chapter 6 details a Bayesian preposterior analysis, extended from the workframe developed in Chapter 4, for the purpose of measurement system design. The analysis will be applied to the aluminium truss bridge presented in the previous chapter. A novel numerical strategy, that reduces the computational effort of the method, comparatively to previous implementations, is also presented. Results are validated against the results of st-id performed in the previous chapter;

Finally, **Chapter 7** highlights the conclusions of the present work and future developments in the field of probabilistic SHM using the developed approach.

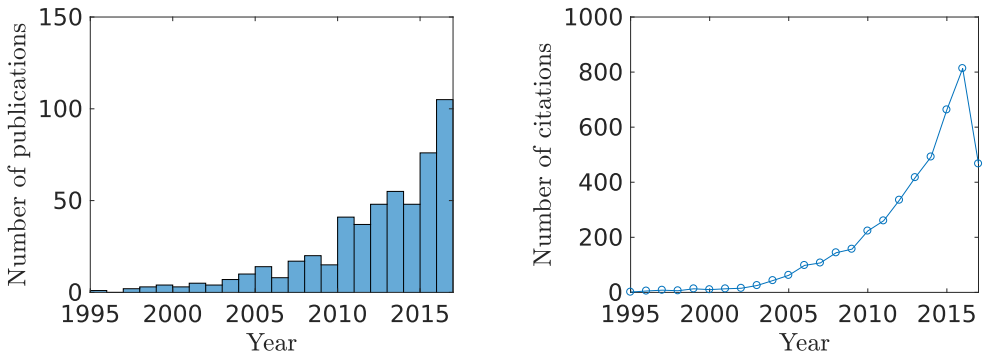
Chapter 2

Literature review

2.1 Introduction

In this chapter a detailed examination of literature in uncertainty assessment of Structural Health Monitoring (SHM) is presented.

By examining the progression of scientific research related with uncertainties and SHM through time, there is a constant increase in the number of publications and number of citations regarding the subject, as can be seen in Fig. 2.1. During a 22 year period 523 publications were reported (source: Web of Knowledge). The "average number of citations per item" of this sample is 8.37, and the "sum of times these publications were cited (without self-citations)", was 3953.



(a) (b)

Figure 2.1: Histogram of scientific research related with: "uncertainties + SHM". Total number of publications (a) and number of times cited (b). Source: Web of Knowledge July 2017.

Obviously, such large amount of research includes several diverse methodolo-

gies, which are applied jointly with different data interpretation mechanisms (data-based or physics-based) to problems with different scales of complexity, e.g. simulated, reduced-scale or full-scale infrastructure.

Therefore, not all of this research will be reviewed in detail, since the current thesis is focused on a hybrid-based interpretation of complex infrastructure. General concepts and relevant probabilistic approaches will also be detailed in order to familiarise the reader with some aspects of the current review.

Thus, the current chapter will be divided into:

- a short introduction to uncertainty quantification in Section 2.2, including types of uncertainties under consideration, mechanisms of interpretation of probability in Sections 2.2.5 and 2.2.6, and statement of the forward and inverse propagation of uncertainty, in Sections 2.2.3 and 2.2.4, respectively;
- an in depth review of Bayesian model updating frameworks, which solve the problem of inverse propagation of uncertainty, detailed in Sections 2.3, 2.4 and 2.5;
- a non-exhaustive literature of data-based Bayesian methodologies, in Section 2.6, highlighting advantages/disadvantages against other physics-based methodologies presented in this review;
- finally, physics-based non-Bayesian probabilistic methodologies will also be presented in Section 2.7, and a summary in Section 2.8.

After examination of the literature in each of these sections, the objectives and relevance of the present thesis will be stated.

2.2 Fundamentals of uncertainty quantification

2.2.1 Model and parameter conceptualisation

Although computational models are so frequently used in a range of scientific fields, e.g. engineering, physics, biology, econometrics, etc. there are always arguments in regards to what is actually represented by a model, its significance and what is its ability to interpret reality.

Beforehand, it is important to establish the concept of a model. As stated by Jaynes [18] a model is a '*mind-projection fallacy*', which represents how the modeller idealises reality. George Box [19] powerful aphorism '*All models are wrong. Some are useful*' corroborates the same idea. In other words, every model will always have

some disparity relatively to the experiment due to assumptions, simplifications, etc. Note also, that in the present thesis a model is viewed in a simplified manner as a black box, which provides a transformation of some input X to some output Y. Its internal operations are not taken into consideration.

Although there is no 'true model', in the sense that it perfectly matches experimental data, that does not mean that there cannot be a 'true parameter', i.e. a parameter set at a value which corresponds to its physical interpretation. Controversially, a true parameter is defined often as one which will provide a perfect fit between simulations/experimental data, see e.g. Beck and Katafygiotis *"a chosen class of structural models does not contain the actual structural system; that is, model error always exists and there are no "true" values of the model parameters"* [20]. In the present thesis, the former definition of a true parameter is adopted.

As an example, if a building is monitored with the aid of an FE model its natural frequencies cannot be matched perfectly with the ones obtained by the calibrated FE model. However, a structural parameter such as the stiffness of the building storey has a reference true value, which depends of the material properties (Young's modulus, Poisson ratio) and of its geometry (second moment of area, area, etc). Finally, it is important to establish that, the value of a parameter can be either assumed as fixed or subject to some inherent variability, e.g. because of some thermal or operational effect. More details on this topic will be detailed in Section 2.2.2.

It is important to mention that the above definition of a parameter presumes that it has a physical interpretation, and is therefore defined as a *calibration parameter*. Additional examples of calibration parameters include the initial strain installed in suspension cables, soil stiffness or other unknown boundary conditions of the structure. On the other hand, *tuning parameters* are considered as parameters which have no meaning in the physical experiment, e.g., mesh density in a finite element simulation or some constant in an empirically postulated material flow law.

A final relevant definition to the current thesis are design variables. Similarly to the model parameters, design variables are inputs of the model which are relevant for a specific type of analysis. For example, temperature fluctuations in a thermal analysis, humidity in a hygrothermal analysis, wind in a fluid dynamic simulation, time in a dynamic analysis, traffic, etc. Note that none of these examples are inherent to the structural system. They depend solely of the particular analysis that the structural system is subjected to. A general perspective of the current discussion is shown in Fig. 2.2.

Having established the definition of model its parameters and design variables, two types of quantification of uncertainty problems can be considered. Either

Figure 2.2: Structural model conceptualisation as a black box. The output $\boldsymbol{Y}$ is a function of the design variables $\boldsymbol{X}$ and the parameters $\boldsymbol{\theta}$.

the uncertainty is propagated through the parameters to the model output, also known as forward propagation of uncertainty, or there is interest in learning the model parameters based on the model output, which is called the inverse propagation of uncertainty problem. Each of these problems will be summarily detailed in Sections 2.2.3 and 2.2.4.

2.2.2 Uncertainties in Structural Health Monitoring

In this section a review of uncertainties that are particularly relevant for hybrid-based data interpretation in SHM shall be detailed. However, it should be stressed that different designations to the ones presented in this section are often used across existent scientific literature. Hence and for the sake of context alternative designations will also be stated.

The current classification will also include which of the two types of effect, *aleatoric* or *epistemic*, is induced by an uncertainty. An aleatoric uncertainty induces errors which have no definite structure or no correlation with each other, whereas an epistemic type of uncertainty occurs when there are unknown dependencies or certain systematic relations between an estimate and the actual value. Up to a certain extent, all real life stochastic processes exhibit these two forms of behaviour. This is important since these two behaviours are directly related with the ability to learn information from data, in terms of precision and/or accuracy. Typically precision/accuracy is mostly affected by the aleatoric/epistemic behaviour of the process, respectively.

Estimation uncertainty: In order to quantify a parameter or a function output, e.g. a stochastic process, with a confidence interval or a a probability density function (PDF) there is always some uncertainty associated with the involved estimation process and associated computations. This is an aleatoric type of uncertainty.

Parameter uncertainty: The assessment of parameter uncertainty can be carried out with probabilistic models, such as a confidence interval or a PDF, and estimation of their statistics. As an example, in Fig. 2.3 it is assumed that a fixed

parameter can be described by a PDF, with associated statistics (mean and variance) which represent the true value of the parameter and its estimation uncertainty, respectively.

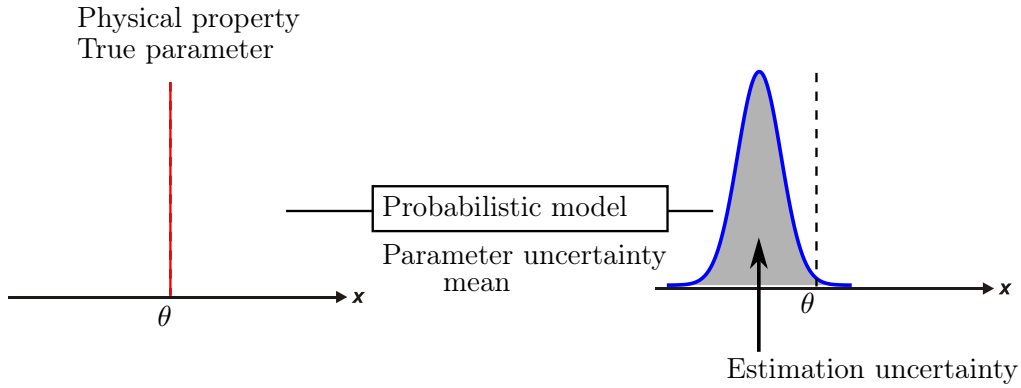


Figure 2.3: Diagram of uncertainties associated with a model parameter which represents a fixed physical property.

Inherent variability: Notwithstanding the above considerations, the physical property which is being represented by a calibration parameter can vary throughout the duration of an experiment i.e. it has an inherent variability (cf. Fig. 2.4). Note that the term "inherent variability" does not refer to a parameter estimation uncertainty but to its variation due to external factors, e.g., changing environmental and/or operational conditions. In contrast with the probabilistic model in Fig. 2.3, estimation of a parameter becomes a much more complex task, requiring large data sets to capture such variability, with increasing complexity for each additional parameter that is considered. Thus, parameter uncertainty and inherent variability are epistemic uncertainties associated with the nominal value and variability of a physical property throughout an experiment.

Model discrepancy: Other than parameter uncertainty, a model based on physics always requires assumptions, simplifications and/or exhibits other forms of inadequacies relatively to the experiment, even if the correct values of its parameters are considered. Model discrepancy is problem dependent, and is also known as model bias, systematic uncertainty, code uncertainty, prediction error, model inadequacy, model form or modelling errors. As an example, Fig. 2.5 displays a simply supported beam and a model of the actual beam. Since, the left support in the actual beam is stiffer than in the model, *all* the displacements of the span are *systematically* different between the model and the actual beam. Additional examples include inaccurate modelling of environmental and operational effects, linear vs nonlinear analyses, number of applied load steps, FE mesh discretisation, etc. Finally, model

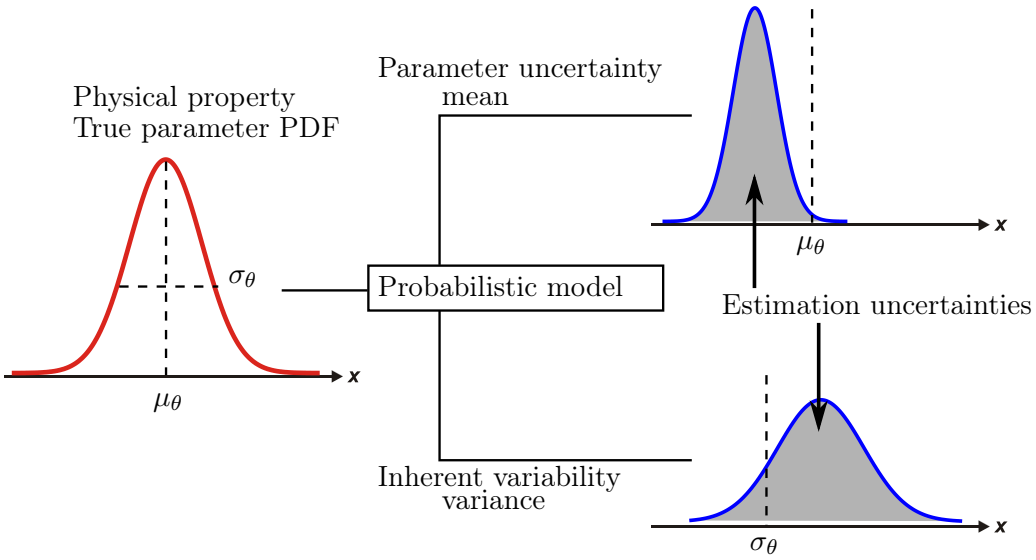


Figure 2.4: Diagram of uncertainties associated with a model parameter which represents a variable physical property.

discrepancy is classed as deliberate or unintended by the modeller, as stated by Legaut et al. [21] and is an epistemic type of uncertainty.

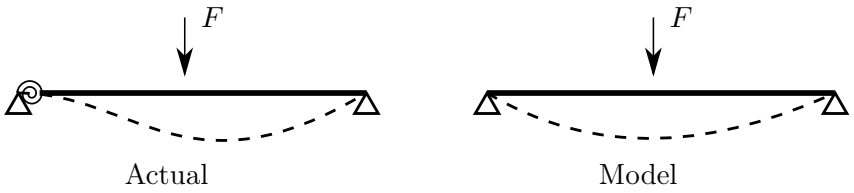


Figure 2.5: Example of model discrepancy in a simply supported beam with a stiff support.

On the other side of the spectrum, **experimental uncertainty**, also known as measurement error, observation error, are errors that exist naturally during the measurement process, even when aiming to reproduce the same experimental conditions from one test to the next. These can develop from several sources, e.g., residual variations of material properties due to plastic or thermal behaviour, or hysteretic cycles, electrical noise, imperfections of sensor devices, etc. Commonly these errors exhibit an aleatoric nature [22–25].

A final source of uncertainty worth mentioning is **interpolation uncertainty** or discretisation error, which occurs when aiming to represent a continuous process in a discrete form. This source of uncertainty would disappear if the model had been run an infinite amount of times, in order to cover an entire model response

surface. This also applies in measured quantities, e.g., when aiming to capture a given continuous signal in a discrete form by selecting a sufficiently high sampling rate.

An illustrative diagram of the sources of uncertainty described above is shown in Fig. 2.6. It shows an instrumented structure from which monitored data is polluted with noise. An FE model of the bridge has been developed to aid explaining available observations. However the modeller is unsure of how certain parameters of the model, which correspond to unknown physical properties of the bridge, have to be fine-tuned. Moreover the simulations are computationally expensive, and therefore only a set of points in a region of interest can be run, raising uncertainty in the intervals between samples. Finally, even if the model parameters were known, the model is limited in its ability to represent reality, and will always have some form of model discrepancy.

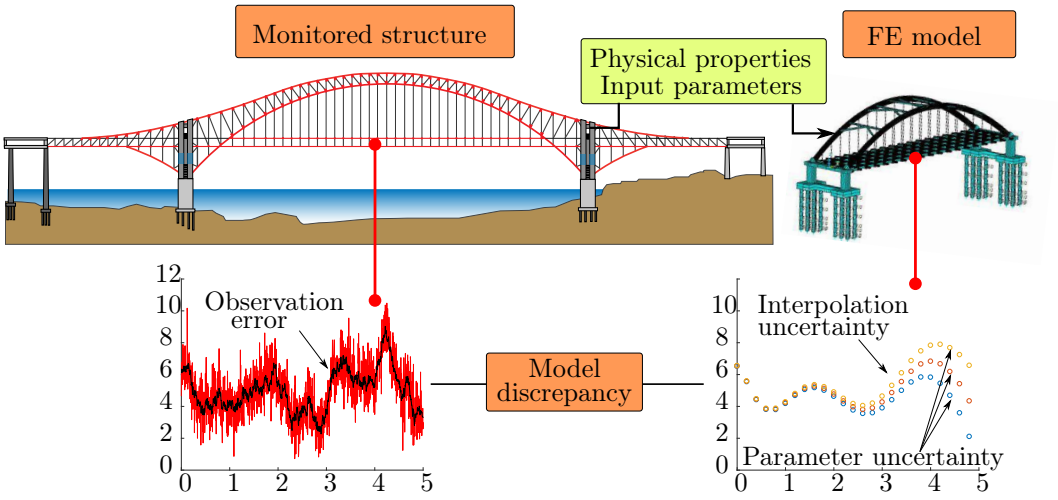


Figure 2.6: Illustration of different sources of uncertainty.

Consequently, a considerable amount of research is dedicated to address these sources of uncertainty, in areas ranging from statistics, engineering, physics, econometrics, applied mathematics, amongst others. As a closing remark, when performing a physics-based data interpretation, the two major sources of uncertainty are parametric uncertainty and model discrepancy.

2.2.3 Forward propagation of uncertainty

Although the present thesis is concerned primarily with inverse quantification of uncertainty, a brief overview of the forward problem will be presented in this section. As mentioned previously, the forward problem concerns the random effect which input

parameters induce on a model output. Common applications which are highly susceptible to this type of propagation are robust optimisation, design of experiments, sensitivity analyses, etc.

In order to investigate such random effects, analysts often perform parametric or sensitivity analyses centred on the variability of input model parameters and the respective output that is propagated across a model. Fig. 2.7 shows a workflow of forward propagation of uncertainty, where tuning, calibration parameters and design variables are propagated through the simulator. The uncertainty in the output is then analysed to estimate how the variability in the input can be apportioned to it.

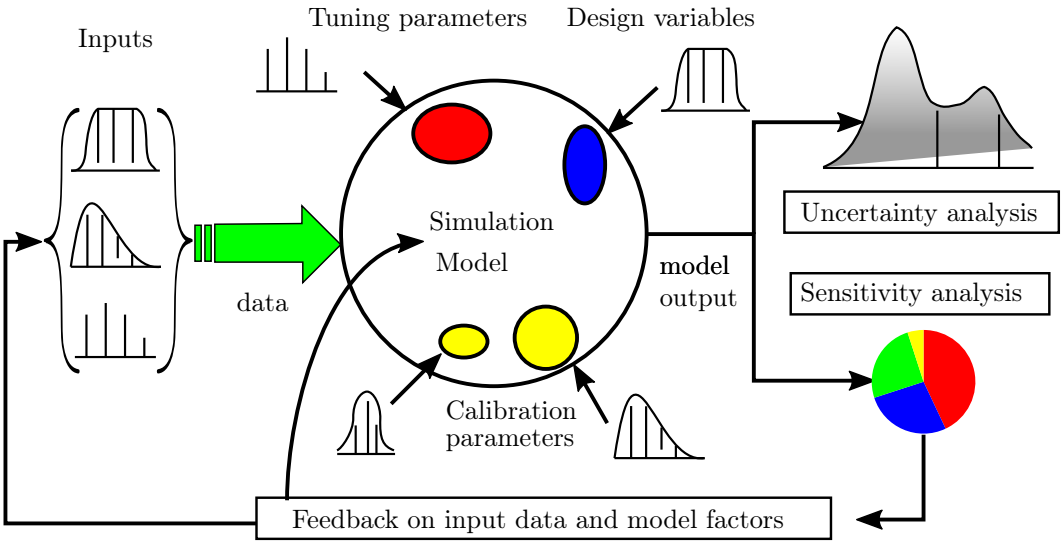


Figure 2.7: Diagram of forward uncertainty quantification, adapted from [26].

To this end, the Monte Carlo methods are commonly used to sample a model response surface and propagate the uncertainties of its input. Unfortunately, when the number of parameters increase and/or a model is expensive to run this approach becomes unfeasible. One possible approach to overcome this obstacle, is to replace the original model with an approximate, faster and simpler model. This type of model is commonly known as surrogate model or emulator. For example, Wan and Ren used a Gaussian process surrogate model to analyse the forward propagation of uncertainty of a full-scale arch bridge FE model [27–29]. The ease of use, low computational effort, and flexibility of this statistical tool considerably reduced the computational effort that would be required to perform Monte Carlo upon the original model. A similar surrogate model approach can be seen in García et al. [30], who used a similar metamodeling approach to study uncertainties in carbon nanotubes

reinforced plates. Two surrogate models, kriging and high dimensional model representations, have been used to replace the original FE model in a multi-scale analysis. Note that kriging [31] and Gaussian processes are extremely similar statistical models in both approach and formulation. A more detailed description will be presented in Chapter 3. Other examples of surrogate modelling techniques include regression models [32], polynomial chaos expansions [33] or radial basis functions [34].

If not through the combination of brute-force analyses with a surrogate model, a factorial analysis can also be used to analyse trends and interactions between factors of a FE model without requiring a large number of runs. Despite not being able to fully cover a model response surface, this analysis provide reasonable insight into the input/output behaviour of model response surfaces. Wiberg et al. [35] and Jesus et al. [36, 37] applied such methodologies to study the effect of key parameters, e.g. ballast nonlinear behaviour, in the dynamic response of railway viaducts. For a more detailed explanation of this class of methods see Box et al. [38].

2.2.4 Inverse propagation of uncertainty

In contrast to the previous section, if it is desirable to know the uncertainty of a parameter which is associated with a model output the inverse uncertainty quantification problem emerges. There are two situations to be considered depending of what is the interest of the analyst

- If the primary goal is to achieve a *best fit* with provided experimental data, then the model parameters can be set freely so that predictions are constrained to match against the data. In this scenario it is irrelevant if the parameters attain their true values or not.

- In contrast, if the objective is to estimate calibration parameters which reflect a performance of interest, e.g. damage level in a certain structural component; or if the deficiencies of a wrong model need to be quantified; or if the model needs to be used to perform predictions over a broader set of input regions; then the parameters have a much greater significance, and should be estimated at their *true values*, irrespective of the fact that output simulations will have a discrepancy against experimental data.

Although several applications in SHM benefit from a "best fit" model calibration, the author believes that most challenging problems can only be improved with a "true parameter" model calibration. The next natural question is "how to estimate a parameter true value?" As shall be mentioned in Section 2.2.7 such task is closely

related with *identifiability*. The remainder of this section will present an overview of formulations associated with the two types of model calibration, highlighting their strengths and weaknesses.

A possible model updating framework can be seen in Fig. 2.8. Essentially, the model and the physical experiments generate simulations D^m and observations D^e, which contain an output variable dataset $\boldsymbol{Y}$, some control variables $\boldsymbol{X}$, and a set of calibration parameters $\boldsymbol{\theta}$. The core methodology has to interpret these datasets, evaluating the uncertainty in the calibration parameters and/or the agreement between the predictions/observations. A validation criterion can be established and checked, e.g. whether model fit is below a certain threshold. If the validation is not conclusive, more data or refinement of the present model is required. Finally, when the validation is successful, the model is considered calibrated and can be used for predictions within the domain where it has been fitted.

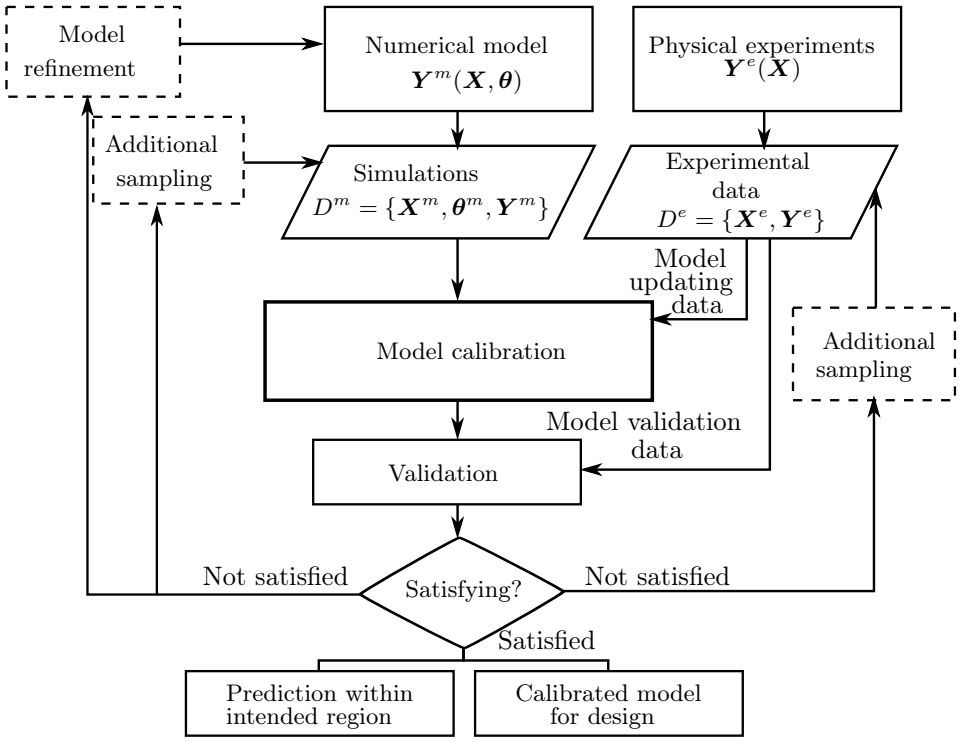


Figure 2.8: Flowchart of the model updating process, from top to bottom: model/ experiments inputs, calibration, validation and additional sampling or prediction, source from [39].

Model calibration formulations

To better present commonly used calibration formulations, let us assume that similarly to Fig. 2.8, simulations from a dedicated physics-based model $y^m(x, \theta)$ and measurements from an experiment $y^e(x)$ are made available. These depend on a number of measurable design variables x and a set of calibration parameters θ.

One of the most commonly used formulations to establish an agreement between these two processes can be stated as

$$y^e(x) = y^m(x, \theta) + \varepsilon \tag{2.1}$$

where ε represents a measurement error which pollutes the measured signal. Despite the measurement error randomness, this formulation has a unique solution for its calibration parameters θ and the experimental error ε. The main criticism to this formulation, is that it does not account for any sort of bias between the predictions and the measurements, in which case the estimated θ will not yield its true physical value.

A second model calibration formulation, which considers an additive function to account for model discrepancy, can be written as

$$y^e(x) = y^m(x) + \delta(x) + \varepsilon \tag{2.2}$$

where $\delta(x)$ represents a discrepancy function which is added to the output simulations y^m. One can think of this term as an error function, which translates the disagreement between observed and simulated data. Also Eq. (2.2) has a unique solution, and can effectively highlight existent bias. However it is unable to learn from physically meaningful calibration parameters and their effect in the model predictions. Illustration of a scenario where a discrepancy function would benefit the calibration process can be seen in Fig. 2.9. The scenario shown in Fig. 2.9(a) would be easily solved using the formulation of Eq. (2.1), whereas the scenario of Fig. 2.9(b) requires a discrepancy function in order to match predictions with the data.

Finally, the most comprehensive formulation which allows to consider all of the presented sources of uncertainty is given by

$$y^e(x) = y^m(x, \theta) + \delta(x) + \varepsilon. \tag{2.3}$$

By solving this equation, firstly, it is possible to learn how adequate the model is by examining the discrepancy function δ. Secondly, the experimental error is also being quantified, translating how well the process is being observed. Thirdly, it is possible

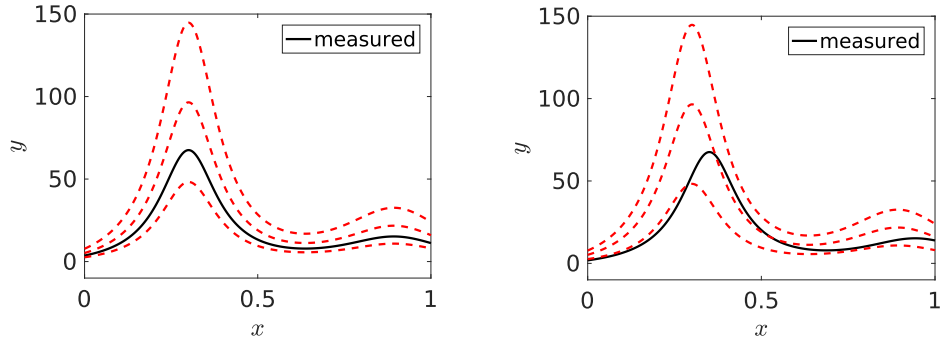


Figure 2.9: Example of several simulations (x,y) for different values of its parameters (dashed lines) against monitored data (solid line) (a), illustrating a scenario with model discrepancy (b).

to learn the true value that the calibration parameters θ take during the experiment.

Although the formulation of Eq. (2.3) allows to assess several uncertainties, at both predictive and experimental level, it is naturally more challenging to solve than Eqs. (2.1) and (2.2). Some of these challenges will be presented in the following sections.

It is also worth mentioning that there are alternative ways of considering the existence of model discrepancy, other than as an additive function. However, the presented form provides a clear separation of each of the considered sources of uncertainty and shall be adopted throughout the present text.

2.2.5 Frequentist framework

The previous sections have established the main definitions associated with data interpretation, associated types of uncertainty and propagation problems. Now it is appropriate to detail some probabilistic frameworks which are commonly used to address such problems.

Frequentist probabilities are a well known and commonly used set of statistical tools, derived from the interpretation of *probability* as a limit of a *relative frequency* in a large number of trials.

Under the umbrella of this interpretation, it is assumed that the convergence limit of a large number of trials $n_t \to \infty$ is the exact true probability of an event x.

$$P(x) = \lim_{n_t \to \infty} = \frac{n_x}{n_t} \tag{2.4}$$

When applying this ideology in order to interpret monitored data with a physics based model it is implied that there is a true model and true model parameters which can be identified.

2.2.6 Bayesian framework

It is appropriate to formally introduce in this section the Bayesian interpretation of probabilities and its fundamental theorem. Several methodologies presented in the following sections are based on this framework. In contrast with the commonly known frequentist interpretation of probability, which represents frequency of occurrence, a Bayesian approach associates *probability* with a state of *knowledge* of an event [18]. This is relevant in order to take into account subjective sources of information e.g. educated guesses, rules of thumb, or other potentially biased judgements.

Bayes' theorem, or the Bayes' rule, is grounded in the notion of conditional probability, which is commonly denoted as $P(A|B)$, i.e. the probability of occurrence of A given or conditioned by B. The vertical bar separates the subset under analysis from information which has occurred. By examining the conditional probability of occurrence of an event A given event B, $P(A|B) = P(A \cap B)/P(B)$, and its inverse, it is trivial to deduce Bayes' theorem. It can be stated simply as

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}. \tag{2.5}$$

and it enables a reverse connection between the conditional probability of two events. This equation, written down by Laplace in 1812 [40], is the most general and commonly known form of the theorem. Interestingly, the theorem was named after Thomas Bayes, given that he introduced a derived and more particular form of the theorem earlier in the 18^{th} century [41]. The several elements of Bayes' rule are often termed as: $P(A|B)$ - posterior distribution, $P(B|A)$ - likelihood function, $P(A)$ - prior distribution and $P(B)$ - marginal likelihood.

The prior represents beliefs upon A which are expressed before evidence is taken into account. Usually the prior is established from past information (a previous posterior PDF) or from elicited experts knowledge [42]. It is worth mentioning that a uniform prior PDF is commonly known as a *non-informative*, *diffuse* or *uninformative* prior, in Bayesian statistics. The simplest type of non-informative prior is one which assigns equal probabilities to all possibilities, also known as the principle of indifference. Bayesian inference which follows such principle and only recovers a maximum point estimate, is equivalent to least squares estimation, simply because no prior belief is being conveyed by the PDF. It should be noted however, that

what is being recovered by Bayesian inference is a *posterior distribution* of parameters, which represents more information than just a point estimate obtained by least squares estimation.

The posterior PDF is the conjugation of the likelihood and posterior distribution. Unfortunately, for most noteworthy applications the posterior distribution is often a multidimensional complex distribution, which requires hundreds of thousands of samples and computational power to determine. Thus, it is not surprising that only in the last decades, with the advent of Markov chain Monte Carlo (MCMC) sampling methods, did Bayesian probabilities became a feasible approach for real world applications.

Finally, the likelihood represents available evidence, in the form of a probabilistic data model. However, it should be noted that the likelihood is not a PDF, and its integration does not yield a probability.

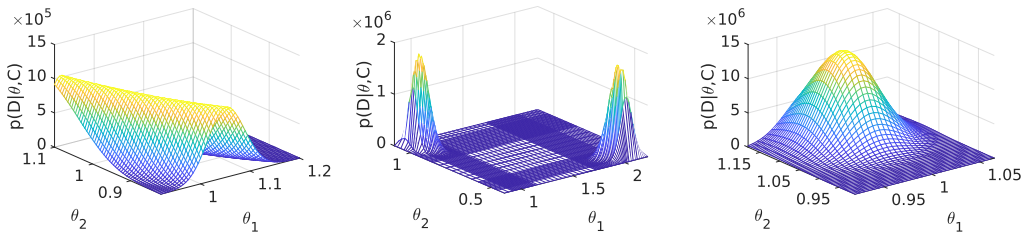
An important closing remark is that the assumption of a parameter as a random variable is often seen as one of the most intrinsic aspects of a Bayesian approach. Note that despite this difference between frequentists and Bayesians, both assume the existence of a true parameter. As stated by Greenland *"It is often said (incorrectly) that 'parameters are treated as fixed by the frequentist but as random by the Bayesian'. For frequentists and Bayesians alike, the value of a parameter may have been fixed from the start or may have been generated from a physically random mechanism. In either case, both suppose it has taken on some fixed value that we would like to know. The Bayesian uses formal probability models to express personal uncertainty about that value. The 'randomness' in these models represents personal uncertainty about the parameter's value; it is not a property of the parameter (although we should hope it accurately reflects properties of the mechanisms that produced the parameter)."* [43].

2.2.7 Model identification and identifiability

Since the inverse problem consists on updating a number of equations, through a relatively small number of calibration parameters the system is generally overdetermined. However the responses which are being used to determine the parameters contain a limited amount of information, dependent of the responses sensitivity against well-chosen parameters. Moreover, with a different calibration parameter every model-instance will always have a discrepancy function that provides a perfect agreement with the measured data. In other words, there is no closed-form solution to solve Eq. (2.3). These factors class the current problem as inverse, overdetermined and ill-posed [44].

Thus, it is important to apply regularisation techniques, e.g. Tikhonov regularisation, and to analyse how capable the model interpretation is to estimate the true values of the model parameters [45–47], i.e. if the model provides good identifiability. Obviously, if the estimated parameters θ attain their true values, the discrepancy function will also highlight the true discrepancy between the model and the experiment, which is an ideal model calibration scenario. The conditions under which identifiability is possible vary, and depend of how observable and stable the monitored data is, how much discrepancy affects the model interpretation, and which formulation is being applied for uncertainty quantification.

As an example, if a Bayesian model is used to identify a spring mass natural frequencies, all of the above factors will affect the shape of the natural frequencies posterior distribution. Identifiability is then classed as **locally identifiable** if the posterior has a region of possible maximum values. If the posterior of a problem attains a global maximum value, it is said to be **globally identifiable**, as defined by Katafygiotis and Beck [48] or Yuen [49]. When identifiability is not possible the model is then said to be **non-identifiable** or unidentifiable. A perspective of these three types of identifiability is shown in Fig. 2.10, which represents the posterior of the two natural frequencies θ_1 and θ_2, given five noisy measurements of one natural frequency Fig 2.10(a), plus two measurements of the second natural frequency in Fig. 2.10(b), and one natural frequency and its mode shape in Fig. 2.10(c). As can be observed, different types of monitored data can significantly improve the problem identifiability.



(a) (b) (c)

Figure 2.10: Posterior distribution for different levels of identifiability of eigenvalues θ_1 and θ_2 from a structural model. Model identifiability illustration, for a (a) non-identifiable, (b) locally identifiable and (c) globally identifiable parameter. Example adapted from [49].

2.2.8 Hypothesis testing

In this section, a few hypothesis testing concepts which are relevant for the present thesis are detailed. The interested reader on hypothesis testing and detection theory can find an excellent reference in Kay's work [50].

The classical problem of binary hypothesis testing assumes that in an observational space X a certain hypothesis, termed the null hypothesis H_o, can be verified by an explanatory statistic and some validation data x. If H_0 is rejected, its dual counterpart, the alternative hypothesis H_1 is believed to be true instead.

The separation of the observational space into the data that supports each of these hypothesis is established by a selected critical value ξ. Fig. 2.11 shows these two regions and the separation point. Associated with the selection of this point two statistical errors become an integral part of the test.

- Type I error: false positive. The null hypothesis H_0 is true, but it is rejected.
- Type II error: false negative. The null hypothesis H_0 is false, but it is not rejected.

The probability of occurrence of errors is denoted as α and β for type I and type II errors, respectively. Note that ξ becomes completely defined, for a particular value of α or β.

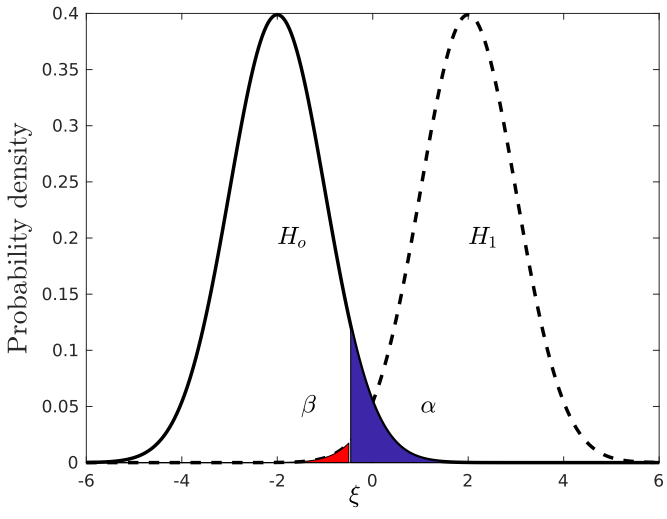


Figure 2.11: Distributions of explanatory statistic under H_0 and H_1, solid and dashed line, respectively. Critical value ξ and incorrect decision regions for type I and type II errors.

Having established these essential concepts, several considerations are left to the end-user. Namely, it is important to select a test statistic that is computationally feasible and that presents a good trade-off between the two possible errors. This trade-off has to be established, since it is not possible to reduce both error probabilities simultaneously. Furthermore, selection of the percentage value associated with errors α or β requires consideration of the risk associated with each type of error. In a Bayesian view, the probability of error P_e which has to be minimised is defined as

$$\begin{aligned} P_e &= P\{\text{decide } H_0, H_1 \text{ true}\} + P\{\text{decide } H_1, H_0 \text{ true}\} \\ P_e &= P(H_0|H_1)P(H_1) + P(H_1|H_0)P(H_0) \end{aligned} \tag{2.6}$$

where $P(H_i)$ is the prior belief of occurrence of event H_i, and $P(H_i|H_j)$ is the conditional probability of deciding H_i when H_j is true.

The minimisation of P_e will depend of which estimator is used, assumptions placed upon the observed data, etc. Although the full proof is not presented here, an example of the condition which a minimum detector would have to follow can be stated as follows

$$\frac{P(x|H_1)}{P(x|H_0)} > \frac{P(H_0)}{P(H_1)} = \xi. \tag{2.7}$$

where the critical value ξ is defined by the prior probabilities of the hypotheses. The equivalent frequentist criteria is known as the Neyman-Pearson theorem.

For example, civil infrastructure owners and decision-makers are usually much more concerned with avoidance of undetected damage, which can lead to catastrophic failure, additional undetected structural damage and other structural misbehaviour, than with costs of unnecessary inspections/replacement of parts. In this context, and having developed a test for validation of the null hypothesis H_0: "there is no damage present on the structure", type I errors are incorrectly flagged occurrences of damage, and type II errors represent undetected occurrences of damage.

2.3 Classical Bayesian framework

Having established in the previous section the fundamental concepts of uncertainty quantification, it is now possible to explore existent Bayesian methodologies in SHM, specifically for inverse uncertainty quantification.

In 1997, Sohn and Law [51] presented the first Bayesian framework for physics-based damage detection. This pioneering methodology estimated a set of non-dimensional parameters $\theta \in [0,\ 1]$, meant to represent damage, by scaling down entries of a stiffness matrix. Its performance has been illustrated with a simulated

analytical example of multistory frames, where modal data has been corrupted with noise, and its sub-structural elements stiffness has been reduced up to 30%. Most of the SHM community cited the completeness of Sohn and Law formulation in regards to the pertinent presence of uncertainties [52], and its ability to classify multiple damage locations.

One year later, Beck and Katafygiotis [20] presented an equivalent Bayesian strategy to the one presented by Sohn and Law. Their framework shall be denominated in this thesis as *classical Bayesian approach or framework*. In theory, this approach could be extended for the design of a measurement system or inference tasks (damage, natural frequencies, boundary conditions, or other structural properties), while accounting for nearly all of the uncertainties mentioned in Section 2.2.2. The main model updating formulation presented by these authors could be written as

$$y^e = y^m(\theta) + \delta + \varepsilon. \tag{2.8}$$

Note the subtle difference between this equation and Eq. (2.3). In Beck and Katafygiotis formulation there are no design variables $\boldsymbol{X}$ considered against the output modal data $\boldsymbol{Y}$, and both experimental uncertainty and model discrepancy have been considered as zero-mean uncorrelated Gaussian.

The basic principle behind this approach is to represent the parameters of a model and other uncertainty sources as random variables with an associated PDF. Afterwards, a joint model class is built from the assembly of these quantities. This process is known as stochastic embedding [53, 54]. From all sources of uncertainty stated in Section 2.2.2, the inherent variability of model parameters was the only one left out. The influence of this work is so considerable, that to this date, early-stage researchers of this field still adopt the classical Bayesian framework.

The original authors of the classical Bayesian framework, Beck and Katafygiotis, sparked several collaborations with other researchers [55–58] in order to consolidate the applicability of their methodologies. Namely Yuen, who addressed all of the previously mentioned topics: model updating [49, 59–62], measurement system design [63] and modal updating [64–66].

2.3.1 Classical measurement system design

Pioneering work on probabilistic MSD is attributed to Papadimitriou, who worked jointly with Beck and Katafygiotis in order to develop a variant of the classical Bayesian framework tailored for MSD [67, 68]. In particular, Papadimitrou et al. [63, 69] established a performance criterion which translates the uncertainty of identified

structural parameters, named information entropy. It can be shown rigorously, that minimisation of the information entropy is analogous to maximisation of the Fisher information [70], and increases the information which a monitoring system obtains about an unknown model parameter. Since the Fisher information is the second derivative of the likelihood, i.e. its curvature, a high value represents a 'more certain' pointed PDF parameter estimate.

Several authors have followed the same approach as Papadimitriou, e.g. Chow et al. [71] and Simoen et al. [72]. Despite encouraging results in simulated and laboratory benchmark examples, no sensor networks of full-scale infrastructure have yet been designed with these methodologies. Another unsolved question was which statistic of the Fisher information ought to be used. The Fisher matrix trace, determinant, etc.

Finally, it must be stressed that MSD is always conditioned by a considerable computational effort. An extensive amount of optimisation routines have been tested to reduce this bottleneck, e.g. polynomial chaos expansion [73], genetic algorithms with parallel processing [71, 74], particle swarm [75], an artificial bee colony algorithm [76] amongst others. Overall genetic algorithms have proven to be well suited for the problem of discrete measurement system design, specially due to their ability to scale well with parallel processing [77, 78].

2.3.2 Bayesian modal identification

The first Bayesian methodology for modal identification was based on the classical Bayesian framework, and it was developed by Au [79], who addressed this topic in both theory and practice [80–82]. The underlying idea behind Au's approach is to calibrate a physics-based model with the same form as the classical forced and damped equation of motion

$$M\ddot{u} + C\dot{u} + Ku = f(t), \tag{2.9}$$

where M, C, and K represent a mass, damping and stiffness matrix, $\ddot{u}$, $\dot{u}$, and u represent the acceleration, velocity and displacement, and $f(t)$ is an excitation force, respectively. Naturally, estimation of the modal properties of a structure (natural frequencies, damping ratios and mode shapes) is the main goal of this procedure.

Numerous practical applications have been tested, e.g. a super-tall building, a 220 m-long Footbridge NF276 across the Tolo Harbor in Hong Kong, and a laboratory four-storied steel frame. Major difficulties arised in scenarios with strong winds, which potentially biased identified modal frequencies. More recently, Au also devel-

oped a rigorous study of the amount of data and sensors required to attain a certain level of precision on the identified modal parameters [83, 84], and validated it in a CityU footbridge in Hong Kong. Finally, Au also addressed the case of modal identification with forced vibration data [85], validating its Bayesian methodology in the UCLA Doris and Louis Factor Health Science Building, and a pedestrian concrete bridge situated in the City University of Hong Kong. This field of research appears to be one of the most successful and well studied examples of Bayesian inference in SHM.

2.3.3 Classical Bayesian damage detection

Naturally, other than measurement system design or modal identification, the classical Bayesian approach was originally designed to address the core topic of SHM, i.e. damage detection in full-scale infrastructure.

Damage detection with the classical Bayesian approach in full-scale structures is still scarce. Some rare examples include a laboratory reduced scale steel bridge by Ntosios et al. [86] and a seven-story full-scale building Simoen et al. [87], shown in Fig.2.12 and 2.13, respectively. In the first example the bridge was damaged in two regions, several damage scenarios have been simulated and used as classifiers of damage location and extent. This procedure used the bridge's modal content as a dataset and a model simplicity criterion in order to select the most reasonable damage explanatory model [88]. For the second example, damage was applied in five stages and the Young's moduli of eight substructures of the building were identified progressively, based on its monitored modal characteristics.

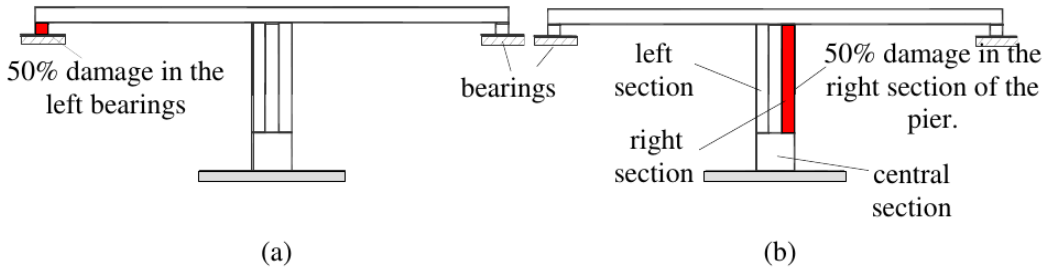


Figure 2.12: Damage features of a reduced-scale steel bridge, source [86].

Results for the first example have shown that a family of three model classes identified the damage, with an error ranging up to 13.2%, i.e. a deviation of 13.2% relatively to the reference values of 50% shown in Fig. 2.12. For the second example the classical Bayesian approach only identified the progressive decay of the Young's

Figure 2.13: 7-story test structure, adapted from [87].

modulus in one out of the eight substructures of the full-scale building. It is recalled that no design variables $\boldsymbol{X}$, as shown in the equations of section 2.2.4, have been considered against the output modal data $\boldsymbol{Y}$ used for calibration. It should also be noted that these examples are relatively well controlled specimens, without significant operational/environmental loads, and the applied damage was significant, scale of 50% or more. Yet the classification of damage ranged from 13.2%, in the reduced-scale bridge, to insensitive in most substructures of the slice building.

In order to explain the poor performance obtained, it is necessary to consider two studies from Moaveni et al. [89, 90] where it has been shown that when using a physics-based approach for damage detection, parameter uncertainty and model discrepancy are the most critical sources of uncertainty which should be taken into account. This conclusion is also valid for model calibration in general and several scientific studies corroborate it. Therefore, it is surprising that a methodology which considered all of these uncertainties explicitly, as the classical Bayesian approach did, had not been as successful when applied to the above examples. The next section will highlight reasons behind this unsolved problem and potential solutions currently under investigation.

2.4 Hierarchical Bayesian model updating

As seen in the previous section, a classical Bayesian framework has been developed and applied in several areas of data interpretation in SHM. In this section a hierarchical Bayesian framework, which is a descendant of the classical framework, will be presented in detail.

2.4.1 Correlation functions in MSD and Bayesian model updating

A first insight into the problem of damage detection with the classical Bayesian approach was established by Papadimitriou and Lombaert [91]. Although the context of their research was MSD, its implications have extended to damage detection, and all other sub-fields of data-interpretation mentioned before.

These authors demonstrated that the behaviour of model discrepancy *was not* uncorrelated as initially assumed, and considering spatial or temporal correlations significantly affected the results of their methodology. They proposed the use of a correlation function to take into account this property, and designed sensor networks which were "spatially aware" of sensor proximity, improving their ability to reduce redundant information from neighbouring sensors. Validation was performed on a continuous beam model, a discrete chain-like stiffness–mass model and a finite element model of a footbridge in Wetteren (Belgium).

There are several possible families of correlation functions. Papadimitriou and Lombaert used a correlation function R defined as follows

$$R(\omega, \Delta) = \exp(-\Delta/\lambda) \tag{2.10}$$

where Δ is the spatial distance between two measurement points, and λ is a positive parameter, known as length scale, which translates how quickly the correlation decays in space or time from $R(\omega, 0) = 1$. The assumed correlation function was used to populate a covariance matrix of model discrepancy, which had non-zero off-diagonal terms, oppositely to a strictly-diagonal covariance matrix in the uncorrelated case.

Despite the improvements brought by this research, other problems have emerged. Namely:

1. it was not clear which correlation function should be used. According with the authors "The correct correlation structure and correlation length over different regions of a structure, however, remain an issue since in the initial experimental design phase no measurements are available to support the selection of a correlation structure consistent with the data and the structural model on

which the optimal sensor location is based.";

2. and the assumed correlation function did not allow for two sensors to be placed at the same location, even if oriented perpendicularly, entailing a loss of relevant information;

In order to improve the first hurdle of this approach, Simoen et al. [54, 72, 92] compared the performance of several correlation functions for Bayesian model updating of a simulated beam case-study. Results suggest that spherical and Gaussian functions yield the best results [72]. Typically, engineers use the Gaussian correlation function because of its differentiability and ease of interpretation. For a visual interpretation of existent correlation functions see Fig. 3.2.

By the same token, Vincenzi and Simonini [93] have brought several contributions to the current topic. Firstly their study included application of different correlation functions, with special emphasis on perpendicularly oriented sensors. Secondly, their work clearly showcased the effect of errors due to spatial correlations, i.e. redundant information because of sensor proximity, from model form errors, i.e. discrepancies between the model and the actual structure (cf. Fig. 2.5). The latter form of error has not been addressed. Finally, validation of their approach was carried out on a simulated simply supported beam with a spring at one end, and in a real five-spans steel bridge, the Correggio footbridge.

Results of designed sensor networks exhibited a strong dependency of model discrepancy due to boundary conditions. Namely, the effective stiffness of a thin steel plate connection of the Correggio footbridge had to be parametrically analysed, in order to diminish model discrepancy. This parameter significantly affected the first and second mode, as seen by comparing the top view of mode 1 and 2 in Fig. 2.14 (a), (b) and (c), (d). In other words, the model had to be updated in order to minimise its discrepancies against the actual structure. It should be noted that the correlation used by these authors is a *spatial* correlation, i.e. it translates the physical proximity between sensors, but the model discrepancy induced by incorrect boundary condition is a *temporal* correlation, which occurs independently of the proximity between sensors.

In short, this section has restated that model discrepancy is a very relevant uncertainty for a physics-based data interpretation, and although considered explicitly in the classical Bayesian framework as a zero-mean uncorrelated Gaussian, those were proved to be strong assumptions. Overall, model discrepancy is problem dependent and extremely difficult to quantify, e.g. Matos et al. [94, 95]. In reality model discrepancy is always correlated, it is not necessarily zero-mean nor Gaussian.

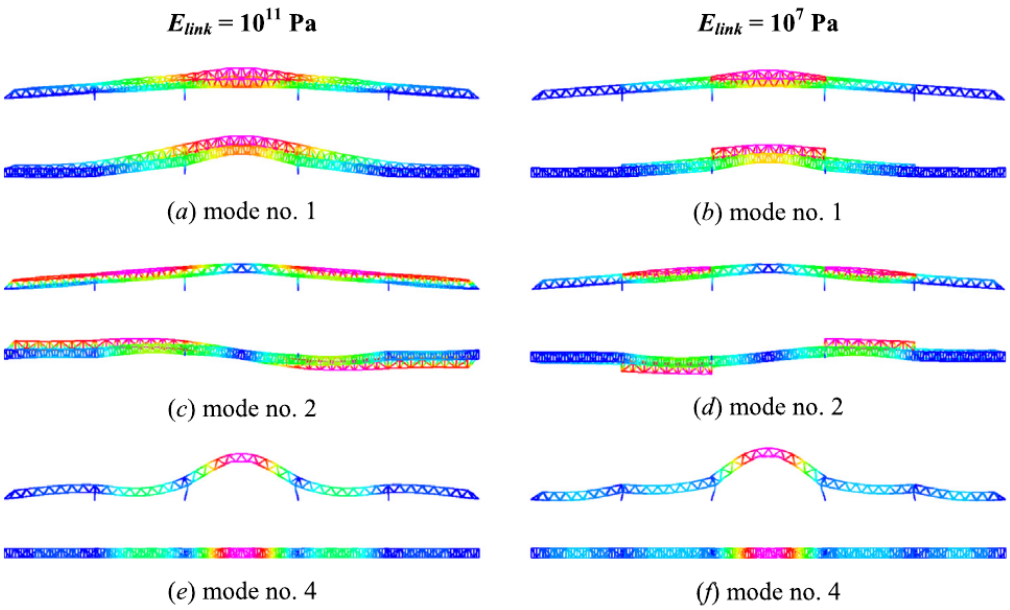


Figure 2.14: Correggio footbridgeplan and elevation of mode shapes no. 1, 2 and 4 obtained with (a, c, e) $E_{link} = 10^{11}$ Pa and $(b, d, f)E_{link} = 10^{7}$ Pa, source [93].

Although formulations which consider spatial correlations do exist, up to this date no temporal correlations have been considered.

Maximum entropy principle

The current section puts in perspective how model discrepancy is considered by the SHM community. As mentioned before, all models have inherent flaws up to a certain extent of their predictive ability, and it is important to acknowledge this fact. Some rare examples which do not consider model discrepancy are visible in Mcfarland and Mahadevan [96] Nishio et al. [97] or Ebrahimian et al. [98]. On the other hand, assuming model discrepancy as a zero-mean uncorrelated Gaussian is extremely common across scientific literature [99–110].

The assumption of a zero-mean uncorrelated Gaussian is justified by the principle of maximum entropy [67, 102, 111], which ensures that by assuming *uncorrelated* errors, the error true value will be captured by the adopted statistical model. This is analogous to assuming that errors are statistically independent, identically distributed and that there is maximum uncertainty. Such assumption is convenient both numerically and as a conservative upper limit. However it automatically negates the possibility of finding patterns and correlations which can be used to assess a model performance or model updating. Researchers such as Ferson [112, 113] argued that

the maximum entropy principle "cannot be justified in real-life problems".

2.4.2 Environmental and operational effects

In addition to the problem of model discrepancy, the classical Bayesian framework presumes that damage is the only causal effect of change in modal data. However this is often not true, due to environmental and operational effects. For example, a temperature increase is known to decrease natural frequencies, in a range of up to 10–20%. Sohn [114] states that temperature changes not only the stiffness of the structure but its boundary conditions.

For the sake of brevity, the term external effects will be adopted throughout the following text to describe operational/environmental effects. In a statistical context these are known as *confounding* factors, since it is necessary to separate their influence from patterns due to damage. This big challenge of SHM had already been pointed out by a considerable number of researchers, such as Doebling et al., Farrar, Sohn et al. etc. [114, 115].

It should be noted that there are numerous sensors which can be used to monitor external effects, such as thermocouples, hygrometers, weight in motion sensors or anemometers. Monitoring data should include at least diurnal/seasonal variations in order to capture the aforementioned effects. Unfortunately, it is not trivial to integrate (or fuse) such data along with modal data when developing and calibrating representative models of a structure. Specific examples for physics-based models are complex thermal effects of structural systems [116], human-structure interaction [117], non-linear contact between a railway and a train wheel [118] and wind pressure on multi-storey buildings [119].

Essentially, the classical Bayesian framework belongs to a class of methods known as vibration-based SHM, which depend solely on vibration data. The new trend of the SHM scientific community is using the structure's output combined with its external effects (temperature, load) as a controllable input. For example, approaches based on temperature effects are known as temperature-based SHM. Examples of methodologies, which consider these sources of variability instead of vibration data for st-id and damage detection, can be seen in Laory et al. [120, 121], Yarnold and Moon [122] and Yarnold and Murphy [123], among others. Although this new trend is more appropriate for long-term monitoring, and provides a more complete data interpretation, it is also more complex and raises an emergent need to quantify uncertainties of physics-based models in order to account for the additional factors.

Summarising the remarks highlighted in the two last sections, the classical

Bayesian approach is limited in its ability to consider model discrepancy and external effects.

2.4.3 Hierarchical Bayesian framework

In order to tackle all of the mentioned hurdles, Behmanesh et al. have developed a *hierarchical Bayesian framework* [124]. As expected, their hierarchical model integrated the previously suggested correlation structure of model discrepancy, and considered external effects through identification of the inherent variability of model parameters.

In short, the underlying idea of this approach is that: external effects can be encapsulated in a structure's material properties; and this interaction can be captured by a hierarchical Bayesian model which parameters vary, according with the structure's external effects. An earlier similar application had already been presented by Ballesteros et al. [125] for uncertainty quantification in structural dynamics.

Oppositely to the classical Bayesian approach, the model parameters of the present approach are expected to vary throughout the monitoring process, and a prior will be set on the descriptive statistics of their PDF's instead of their values. It equates to establishing a prior of a prior (commonly termed an hyperprior) in a hierarchical multi-layered structure.

As an example, suppose that the Young's modulus $\Theta = E$ of a structure has to be identified based on observed modal data Y, which is influenced by temperature or loading conditions, cf. Fig. 2.15. A hierarchical model, which can reproduce the same behaviour as shown in Fig 2.15 through variation of its parameter, can then be defined as follows

$$\begin{aligned} Y &\sim p(y|\Theta) \\ \Theta &\sim p(\theta|\Phi) \\ \Phi &\sim p(\phi), \end{aligned} \tag{2.11}$$

where "$\sim$" means "is distributed according to", the PDF $p(\theta|\Phi)$ that describes the Young's modulus is unknown, and depends of a set of parameters (often named hyperparameters) ϕ, which have to be estimated from the prior $p(\phi)$ and the observed data.

Also in the hierarchichal Bayesian framework there is an assemblage of a global joint posterior PDF of model parameters, for each of the considered hierarchic levels. It is clear that the Young's modulus parameter benefits from considering its inherent variability due to temperature effects, and its identification will result in a bimodal distribution, as can be seen in Fig. 2.16, where higher/lower values correspond to higher/lower monitored frequencies, respectively.

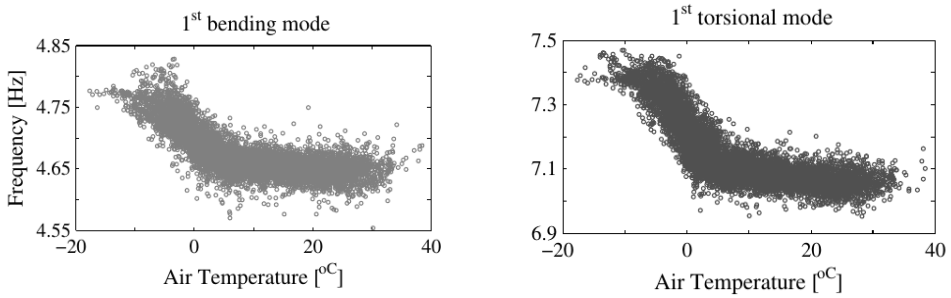


Figure 2.15: Effects of ambient temperature on identified bending and torsional natural frequencies [126, 127].

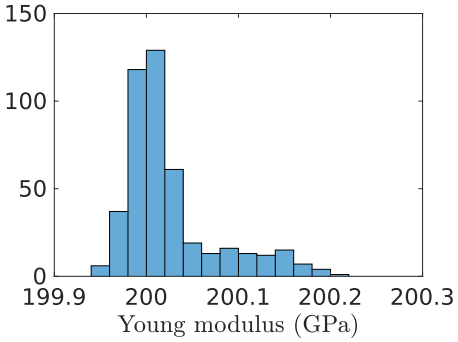


Figure 2.16: Bimodal distribution of Young's modulus associated with modal data of Fig 2.15.

The main advantage of this framework include its extensibility, since it is possible to specify as many hierarchical levels as needed, and its ability to unveil the inherent variability of identified model parameters. On the downside, the hierarchical framework requires identification of two statistics for each calibration parameter, larger data sets, and naturally, its posterior is more complex and difficult to estimate.

The first validation of the hierarchical Bayesian framework was carried out in a three story simulated frame, whereby the storeys stiffness were identified from the frame's natural frequencies, which varied according with prescribed PDFs. A case-study has been considered to access its ability to address model discrepancy, namely a frame with a flexible foundation (cf. Fig. 2.17). Results indicate that the approach accurately captures external effects, but under the presence of model discrepancy the identified stiffness of the first story has a 20% relative error.

Behmanesh work suggests that model discrepancy is still not as comprehensively detailed by the hierarchical Bayesian framework as it would be required to perform model updating or st-id in SHM, and reinforces that this source of uncer-

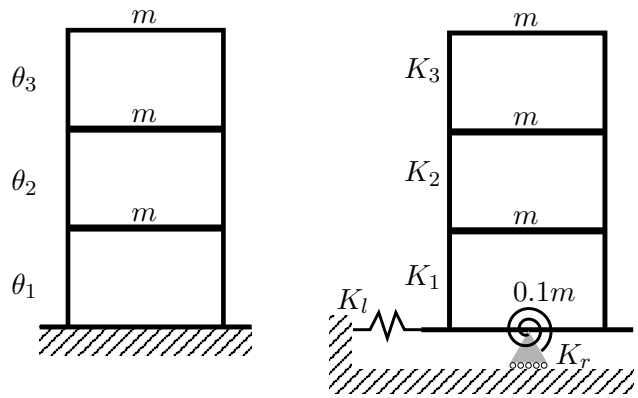


Figure 2.17: Three story shear building example with model discrepancy. Idealised model and its parameters θ_i (left) and actual frame with flexible foundation and stiffness of each floor K_i (right), for floor $i = 1, 2, 3$.

tainty should be considered with very weak assumptions [128, 129].

2.4.4 Hierarchical structural identification and damage detection

As we have seen, external effects are encapsulated in the inherent variability of model parameters, as demonstrated by the identified Young's modulus in Fig 2.16. Hence, if the PDF of any identified property shifts abruptly, it could easily be distinguished from the variability of other external factors, accurately indicating the presence of damage. However this assumption presumes that:

1. the parameters have been accurately estimated. An illustrative example would be the estimation of the cross-sectional area of a beam for the purpose of assessing its reserve capacity;

2. and the parameters influence on the model output must be sufficient to capture all types of local external effects.

In this manner, Behmanesh et al. have applied the hierarchical Bayesian framework to a Dowling Hall footbridge on the Tufts university campus [126, 130, 131] and a nine-story building [132]. Damage was induced by a set of concrete blocks loaded on the footbridge deck for three days. The Young's modulus of the bridge (cf. Fig. 2.18(a), 2.18(b)) and its variations due to temperature and loading were selected as model parameters to identify, subsequently testing the model for damage detection. Conclusions report that several simplifications are required in order to successfully apply the proposed framework for damage detection. All monitored data was recorded in warm temperatures and some of the bridge mode shapes have been

(a) (b)

Figure 2.18: Tufts bridge: photo (a) and FE model (b).

excluded from the identification process. However, the most remarkable aspect is that an uncorrelated zero-mean Gaussian has been used, once again, to represent model discrepancy. It is not certain why did the authors revert to the same model discrepancy form as the classical Bayesian approach, but in the author's opinion, computational aspects played a major role in this decision. Deriving and solving the analytical solution of the updated damage parameters (added mass) also poses a complex formulation and additional numerical challenges.

More generally speaking, three fundamental problems might be thought as to why comprehensive Bayesian methodologies, such as the classical and hierarchic frameworks, are not widely applied in SHM practice.

- When these formulations are applied to st-id or damage detection, there are several parameter estimation problems, arguably due to the presence of large model discrepancy and external effects. The hierarchical approach is however, able to cope with the variability of these parameters due to external effects.

- Both formulations are unable to efficiently update physics-based models, since the information of the model discrepancy patterns and correlations is not being estimated. Instead this uncertainty is estimated as a zero-mean uncorrelated Gaussian. Some progress has been made in this regard with the hierarchical approach, by considering a correlation function. However, all practical case-studies have neglected this feature.

- Due to their comprehensiveness, both frameworks require a considerable computational effort which limits their application to large-scale infrastructure. As stated by Zhang et al. [133], this computational hurdle increases with the number of model parameters and the number of times that the FE model has to be run. It is remarkable that the hierarchical approach has been applied to

several case-studies, despite prone to having more parameters to update than the classical approach.

2.5 Modular Bayesian model calibration

2.5.1 Modelling assumptions

Interestingly, another Bayesian methodology was developed almost in parallel to the work mentioned in the previous sections. It was developed originally in the statistics community, as a milestone work from Kennedy and O'Hagan [134, 135], and it has been named as modular Bayesian approach (MBA). It dates just two years after the original classical framework from Beck and Katafygiotis. This fact, coupled with the different disciplinary field in which it was developed, might explain why it has not received as much attention in the SHM community. Although the MBA aims at solving the same fundamental problem of uncertainty quantification and model calibration as stated before, it uses a different approach and assumptions.

Firstly, it is important to notice that the main equation of the MBA is formulated differently, see Eq. (2.3), from the classical and hierarchical framework, Eq. (2.8). Although the same sources of uncertainty are present, the MBA considers the design variables x. Hence, it assumes that the observed (and modelled) output are *dependent* of a variable which can represent external effects, e.g., temperature, wind, traffic load, etc. In contrast, the hierarchical Bayesian estimates and encapsulates this information in the parameters of its hierarchical model.

Secondly, the MBA uses Gaussian processes to represent the computer model and model discrepancy. A Gaussian process is a statistical model which assumes for simplicity, that the underlying process that it portrays is Gaussian. It generalises the formulation of a Gaussian random variable to infinitely many variables, also known as a multivariate normal distribution. A more rigorous explanation will be presented in Chapter 3. As shall be detailed below, this choice of representation provides several advantages.

Finally, it should be mentioned that up to this point only the assumptions placed upon the uncertainties of Eq. (2.3) have been discussed, but assumptions on how these uncertainties *interact* and are *approximated* also significantly impact the results. For example, considering that parameter uncertainty, model discrepancy and observation error are statistically independent eases several computational aspects, but it worsens how well these uncertainties are approximated. An alternative designation for this particularity of the the MBA is "neglecting the second-order effects of uncertainties". Such aspect is a trade-off which the authors of the MBA

advocate, and stands oppositely to the classical or hierarchical Bayesian ideology, typically known as fully Bayesian approaches.

2.5.2 Advantages/disadvantages

A consequence of the remarks described in section 2.5.1, is that a user of the MBA needs to consider external effects explicitly when developing a physics-based model of a structure, whereas a user of the hierarchical Bayesian framework does not. Obviously this requires more work and leads to additional model discrepancy (since it is not trivial to simulate external effects), but in principle, and since the discrepancy function also depends on the design variables x, it is possible to account for existent biases and update the model accordingly. It is arguable which is the most appropriate way of considering external effects, and it would be interesting to compare the performance of the MBA against the hierarchical Bayesian framework in this regard.

Additionally, the MBA has weaker assumptions in relation to model discrepancy, since the Gaussian process which portrays this source of uncertainty is not zero mean or uncorrelated. For illustrative purposes, the formulation of model discrepancy of the classical, hierarchical and modular Bayesian frameworks can be stated as follows

Classical Bayesian	Hierarchical Bayesian	Modular Bayesian
$\delta \sim \mathcal{N}(\mathbf{O}, \boldsymbol{V})$	$\delta \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{V})$	$\delta(x) \sim \mathrm{Gp}(\boldsymbol{\mu}(x), \boldsymbol{V}(x, x'))$
$\boldsymbol{V} = \sigma_{ii}^2 I$	$\boldsymbol{V} = \sigma_{ij}^2$	$\boldsymbol{V}(x, x') = \boldsymbol{\Sigma}^2 \otimes \boldsymbol{R}(x, x')$

where $\boldsymbol{\mu}$, $\boldsymbol{R}$, σ_{ij} and $\boldsymbol{V}$ represent mean vectors, a correlation matrix, covariance terms and the matrices which contain these terms, respectively. Note that the MBA form is the only one which depends explicitly of the design variables x.

Moreover, by not considering uncertainties second-order effects, the MBA has a lower computational effort than the classical and hierarchical Bayesian approach. Note that implementation of the MBA complete framework is quite demanding and requires extensive use of numerical integration techniques such as quadrature or MCMC integration.

The main disadvantage of the MBA lies on its difficulty in identifying the true model parameters. It is a situation analogous to Fig. 2.10(a), where there are several optimal values of the posterior distribution for the same equation of model calibration. This is a direct consequence of not considering uncertainties fully, and has been the subject of criticism of the MBA in the statistics community. For example, Mcfarland and Mahadevan [96] and Ling et al. [136] attempted to correctly estimate the true model parameters/model discrepancy with the modular Bayesian

approach, but were not able to, since the terms of Eq. (2.3) were highly correlated with each other, biasing the results. It is recalled, that certain areas of research such as damage detection and reliability analyses often require that estimated model parameters attain a physically meaningful value.

Another important remark is that, the MBA does not allow to consider the inherent variability of model parameters, at least not as explicitly as the hierarchical Bayesian approach. To achieve the same result, it would be necessary to consider the actual statistics of a model parameter (mean and covariance) as parameters to identify, and run a much larger set of simulations with different moments.

Despite these criticisms, several authors focused on the ability of the approach to accurately predict physical processes and its popularity remained high, as illustrated in works by Higdon et al. [137, 138], Bayarri et al. [139] and Ranjan et al. [140]. Higdon characterised material properties of steel cylinders during an implosion by a high explosive with the MBA, using highly multivariate data converted through principal component analysis [141]. Multiple parameters were considered and their posterior distribution was sampled via MCMC, namely the Metropolis–Hastings algorithm [142–145].

2.5.3 Identifiability of the modular Bayesian approach

The interest in the MBA amongst engineers and researchers has recently reemerged, due to Arendt et al. These authors set a new milestone by analysing the subject of identifiability [45, 46], how it affected the MBA, and how it could be improved. According to their first insightful paper [39], the original approach formulated by Kennedy and O'Hagan considered only a single response of measured/simulated data. They proved that, unless under some specific conditions, the single response case will fail to identify the true structural parameters, and proposed instead a multiple response approach [146], which allows for a more informative model and better identifiability. Validation was carried out in a simulated example of a simply supported beam subject to a mid-span point load. No observation error was considered and only a single structural parameter was calibrated. Moreover, no ambient effects, e.g. temperature, wind, have been considered by these authors.

In a work by Jesus et al. [222], which is included in this thesis: the practical feasibility of the MBA has been shown in a laboratory reduced-scale aluminium truss bridge subject to thermal expansion, illustrating the ability of the enhanced MBA proposed by Arendt in practice, highlighting the relevance of diversely natured responses, and how it benefits the identifiability of the methodology.

2.5.4 Modular Bayesian measurement system design

The topic of MSD which improves the ability of the MBA to learn true parameters has been studied by Arendt et al. [147] and Jiang et al. [148]. These authors used identifiability as a performance criterion to select most informative responses. Although the authors have reached the conclusion that "responses which have a diversified nature usually provide better identifiability", such conclusion is still relatively vague, and does not explain what is the cause for a response which provides good identifiability. Also, similarly to the work mentioned in the previous sections, a major drawback is the computational effort required for optimisation of the measurement system.

2.6 Data-based Bayesian approaches

In this section Bayesian methodologies under data-driven approaches, i.e. methods which interpret data without resorting to a physics based model, are briefly reviewed.

Figueiredo et al. [149, 150] research focuses in assembling a cluster of relevant features/observations of a structural system based on a Bayesian approach, in order to establish a healthy-state baseline. Subsequently, a decision making criterion (the Mahalanobis squared-distance) was used to classify the outliers which represent damage or other sources of variability. It should be noted, that although Figueiredo's methodology is able to individualise a cluster containing damage features, it is unable to classify it differently from other sources of variability, being prone to the occurrence of Type I errors (see Section 2.2.8 for a definition). Validation was carried out with the Z24-bridge dataset.

As seen above, the problem of signal separation between damage features and external effects is possibly the most sought-after feature in data-based approaches. Furthermore, occurring damage can manifest itself through a discontinuous divergence from a smooth and continuous structural behaviour. Thus, unusual patterns such as outliers could indicate the presence of damage. In this context, Worden and Cross [151] and Becker and Worden [152] developed a Bayesian regression tree Gaussian process, able to fit itself to non-smooth trends of a structure's behaviour. thus accounting for sudden changes in the structural behaviour, potentially caused by damage. It should be noted that inherent smoothness and continuity are a basic assumption of Gaussian processes, hence the novelty of the approach. Major difficulties were observed when emulating certain complex trends of a duffing oscillator response.

A successful case of damage prognosis of structural components subject to

fatigue can be seen in a work by Gobbato et al. [153], by using a recursive Bayesian method. At a component level, without any external influences, and with a large database, this approach provided optimal results, similarly to the most successful application of SHM, i.e. rotary systems. However, in real-life problems of structures sensitive to fatigue, such as bridges, these conditions rarely hold.

Finally, examples of linear regression models updated through Bayesian inference can be seen in works by Enright et al. [154], and Arroyo et al. [155]. Results from these works demonstrate a typical overfitting problem associated with the solution of inverse uncertainty quantification problems. It is believed, that most of current research do not focus as much towards this type of approach.

2.7 Physics-based uncertainty quantification methodologies

It is reasonable to state that research using Bayesian probabilities are a large fraction of uncertainty quantification in SHM. Other physics-based methods related to the quantification of uncertainties in SHM include fuzzy numbers, Kalman filters, sampling methods, model falsification and Markov processes.

2.7.1 Fuzzy numbers

Research of physics-based data-interpretation using fuzzy numbers logic in SHM is very scarce, with only a few authors involved. Namely, Erdogan et al. [156, 157] who has developed a fuzzy steepest descend algorithm for model update and damage identification, having applied it in benchmark grid structure of the University of Central Florida. It should be noted that non-parametric uncertainties (such as model discrepancy) were not considered. Other authors such as Chen [158] and Meyyappan [159] developed methodologies under the same scope, the latter having validated his approach on a steel bridge located in Missouri. Finally, Carden and Brownjohn [160] used a combination of Stochastic Subspace Identification and fuzzy numbers clustering for damage detection in the Z24 bridge and a tall building continuosuly monitored for several years. Results show that the method identified damage successfully, even though it did not separate changes from temperature and damage. Simply put, the existent damage features were so severe that their presence became noticeable, despite the masking effect of temperature.

2.7.2 Kalman filters

In contrast, the literature in the subject of Kalman filters and their application to SHM is extensive and includes several applications to large-scale infrastructure. In Fig 2.19 a diagram of the basic principles of Kalman filters is shown.

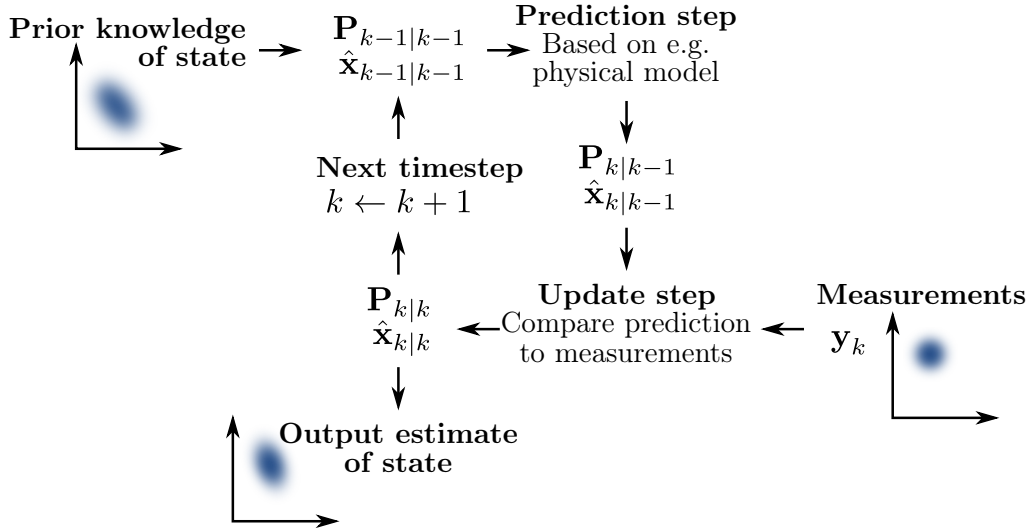


Figure 2.19: Basic working principle of Kalman filter, adapted from [161].

One common superclass is the extended Kalman filter (EKF), which is the nonlinear variant of these methods, having been used extensively in several industrial applications, such as telecommunications, Global Positioning System (GPS), etc. Ebrahimian et al. [162], applied the EKF in a cantilever steel bridge column for identification of three material parameters, and identification of six unknown material parameters of a three-story three-bay moment resisting steel frame. Advanced nonlinear constitutive models were used to accurately describe the dynamic behaviour of such examples [163]. Results have shown an excellent performance against observation error, but a limited estimation in the presence of model discrepancy. Such conclusions lend support to the fact, that even a sophisticated structural model eventually reaches a predictive limit, and is preferable to quantify its uncertainties, rather than persisting on its update. Finally, more recent work from the same authors [98] include a batch Bayesian approach, with a similar workflow as the classical Bayesian framework (discussed above).

Other authors such as Jin et al. [164, 165], Yang et al. [166] and Ding and Guo [167] also applied this superclass of Kalman filters, despite limited to simulated case-studies. Capellari et al. is another author that first used a hybrid particle Kalman filter subclass for damage detection [168] and subsequently shifted its re-

search focus to Bayesian probabilities.

A second Kalman filter class worth mentioning is the ensemble Kalman filter by Ghanem and Ferro [169]. This methodology has been applied for damage detection under the presence of model discrepancy and strongly non-Gaussian signals. Despite providing some advantages relatively to the EKF, which always rely on the Gaussian assumption, this method also did not consider the systematic uncertainty of model discrepancy. More tests with full-scale case-studies are required for validation purposes. A final relevant work, developed with the goal of overcoming identification divergences, large computational effort and storage capacities required by the traditional Kalman filters, is the two-step Kalman filter by Lei et al. [170]. This algorithm can be implemented in micro-processors for damage detection and structural identification and retains all the advantages of the traditional Kalman filter. However also this methodology requires further validations for its feasibility with full-scale structures under environmental/operational conditions.

2.7.3 Model falsification

Another interesting methodology by Goulet and Smith [171, 172] is the probabilistic model falsification. It applies the concept of hypothesis testing in order to refuse or accept a model instance, a concept established formally by Popper [173]. It works by establishing a hyperlatin grid space with model instances, which are then classed according to a set of validation criteria. It does consider model discrepancy as a systematic error, and does not assume it to have a Gaussian structure, unlike most of the previous works. The approach of the method is different, in the sense that the objective is not to identify an optimal model instance, but rather to falsify the ones that do not explain the physics of the problem. Parametric uncertainty is represented by a uniform distribution, which unfortunately does not provide a considerable insight into how these values might change. Despite this, the methodology considers identifiability as a performance metric to determine an optimal sensor configuration [174]. Its validation can be seen in full-scale case-study of Tamar bridge (UK) and Langensand Bridge (Switzerland) [128, 175]. It should be noted that this methodology returned parameter confidence intervals that are physically unrealistic, such as the stiffness of springs in Tamar bridge embankments, at $k = 10 \times 10^{11}$ kN/mm, seen in [120, 176].

2.7.4 Sampling methods

In regards to sampling methods, the amount of research is not very extensive, most likely due to the high computational effort that is required to represent the response surface with a dedicated FE model. For example, Fang and Pereira [177, 178] used a factorial experiment response surface method for damage detection. Model discrepancy has been assessed by the R^2 criterion [179], which translates how much variability there is, which can be explained solely by a model. This means that the applicability of this methodology is limited to the agreement between predictions and measurements, i.e. modelling expertise. This methodology has been validated in a simulated reinforced concrete (RC) beam, a reduced scale RC frame, and the full-scale I-40 bridge [180]. Other recurrent research areas, where often model discrepancy is considered as a zero-mean Gaussian, include least squares as seen in Huang et al. [181], stochastic search methods by Stull et al. [110], and finally, a MCMC sampling by Li et al. [182].

2.8 Summary

This section summarises the information presented in the chapter and contextualises it in the framework of this thesis. Several probabilistic model updating methodologies have been developed through time for several purposes, namely: structural identification, measurement system design, operational modal analysis and damage detection. Particular emphasis was given to the classical, hierarchical and modular Bayesian model updating frameworks, which have been rigorously detailed. A small section has been devoted to Bayesian data-based methodologies. Finally, an overview and shortcomings of other physics-based probabilistic methodologies has been highlighted, including Kalman filters, model falsification, fuzzy numbers logic and sampling methods.

Table 2.1 presents a summary of the discussed Bayesian methodologies, features and advantages/disadvantages. As can be seen, the classical Bayesian approach presents the worst case scenario, with a high computational effort, strong assumptions of model discrepancy and do not considers environmental/operational effects. Conversely, the MBA has a low computational effort, due to the degree of approximation of its uncertainties, a model discrepancy with weak assumptions, and its identifiability depends of considering multiple responses. However both of these methodologies do not explicitly consider the inherent variability of model parameters comparatively to the hierarchical Bayesian framework, whose assumptions regarding model discrepancy are in an interesting mid-term between the modular and classical

Bayesian approach.

	Classical	**Hierarchical**	**Modular**
Inherent variability of model parameters	not explict	Yes	not explict
Model discrepancy (assumptions)	strong	mild	weak
Computational effort	high	high	low
Uncertainties order of approximation	high	high	first

Table 2.1: Summary of presented Bayesian methodologies and their advantages/ disadvantages.

There is no shortage of agreement within the research community, in that model discrepancy and parameter uncertainty are the most severe uncertainties for model updating. However, most of the present research in Bayesian methodologies either does not consider model discrepancy, or considers it with very strong assumptions, such as a zero-mean uncorrelated Gaussian behaviour. Multiple methodologies seen throughout this review follow the same path, under the umbrella of the maximum entropy principle. In the author's opinion, considering model discrepancy in such way negates the possibility of finding patterns and correlations in this source of uncertainty, which are vital to assess the performance of the predictive model and guide its updating process.

Regarding non-Bayesian physics-based probabilistic methodologies, literature suggests that Kalman filters have the overhead interest of the research community. Several classes of Kalman filters are being developed, with a two step Kalman filter algorithm from Lei et al. [170] an interesting option that aims to surpass the limitations of the traditional Kalman filter. Fuzzy numbers logic from Erdogan and the modal falsification method from Goulet are also interesting methodologies, which despite concerning identifiability issues, do show capability to comprehensively quantify uncertainties from several different sources. It must be noted that validation of these methodologies in full-scale structures always requires costly tests or inspections and is very rarely seen. Remarkably, with Bayesian methodologies this is specially notorious, given the amount of research already developed.

Finally, the present review suggests that the MBA accounts better for model discrepancy than the classical or hierarchical Bayesian frameworks. Therefore, the same concept as the original MBA formulation shall be adopted in this thesis. It is necessary to test how the MBA considers environmental/operational effects explicitly, comparatively to the hierarchical Bayesian framework. Moreover, the current implementation of the multiple response MBA has not yet been tested in practice, nor do allow for multiple model parameters calibration.

Thus, there are six specific issues which shall be addressed.

- The enhanced MBA needs to be able to identify multiple model parameters. A routine based on the Metropolis Hastings algorithm has been implemented for such task and will be detailed in Chapter 4, Section 4.6.2.

- Environmental and operational effects should be explicitly considered in the design variables of the methodology, showcasing its ability to cope with these effects. Practical illustrations are detailed in Chapter 5.

- Naturally, the estimated discrepancy function has to cope with inadequate modelling of external effects and/or structural behaviour e.g. boundary conditions.

- On the basis of the enhanced MBA, a measurement system design methodology shall also be developed, in order to select the responses which allow the most accurate parameter estimation. Chapter 6 highlights an optimal sensor configuration analysis for the enhanced MBA and its practical application for SHM.

- A first practical validation of these methodologies shall be carried out in a reduced-scale aluminium bridge and the Tamar long suspension bridge.

- Finally, computational aspects assume a relevant importance for the present work. Although the computational effort of the MBA is low when compared with its fully Bayesian counterparts, its numerical stability is equally important. To this end, a numerical factorisation treatment is to be applied to the multiple response Gaussian processes formulation. Section 3.7 of Chapter 3 provides further details of this enhancement. This is relevant since, even for Arendt's simply supported beam example [146], numerical instabilities have been reported.

Chapter 3

Single and multiple response Gaussian processes

3.1 Introduction

In this Chapter the mathematical formulation of a single and multiple response Gaussian process (mrGp) surrogate model, that fits a multi-input–output function is presented. An SHM practitioner who intends to use the MBA should become familiar with the contents of this chapter before moving to Chapter 4. The single and multiple response case will be detailed sequentially in each of the chapter sections, in order to identify the similarities and differences between each formulation.

Most of the material presented hereafter is classic in the statistical data analysis literature, and is inspired by Rasmussen and Williams [183], Lophaven, Nielsen and Søndergaard [184] and O'Hagan et al. [185–188]. Section 3.7 presents the thesis author's enhancement of the mrGp formulation, which allows to improve the numerical stability of the mrGp likelihood function. This is a core novelty of the present work.

Before outlining the contents of this chapter, a short description of a Gaussian process is given. Let us assume that a function of interest $y(x)$ can be approximated as a spatial random process, which is assumed Gaussian for simplicity. The function $y(x)$ is commonly known as a *latent function*, and its input x can be considered as design variables, calibration or tuning parameters. In the current text it is irrelevant which of these three types of input is considered. Hence

$$y(\cdot) \sim \mathcal{N}(m(\cdot), V(\cdot,\cdot))$$

where $\mathcal{N}$ represents a normal distribution, with mean and covariance functions m and

V, respectively. These functions in turn depend of some specific parameters, which shall hereby be denoted as hyperparameters $\boldsymbol{\phi}$, in order to distinguish them from calibration or tuning parameters of a physics-based model. Usually two situations are relevant to the analyst, namely:

- having observed the latent function, a training dataset is made available and needs to be used in order to estimate the hyperparameters. A Gaussian process approximated in such a way is said to be fitted, trained or regressed to the data;

- the Gaussian process can also generate predictions of the underlying latent function, provided that it has been trained and the hyperparameters have been estimated. This process is also commonly known as emulation.

It is also possible to perform a probabilistic classification of new data, in the sense of how likely it belongs to the latent function. However classification using Gaussian processes will not be detailed in the present work. As a final remark, it should be noted that the above described procedure is analogous to Bayesian inference, where a set of functions (priors) are updated based on available data (likelihood) to obtain a trained model (posterior).

Based on the above mentioned description the present chapter will be divided in five points.

- In Sections 3.2 and 3.3 mean and covariance functions priors, which have been assumed for the current work, and associated hyperparameters shall be carefully detailed.

- Section 3.4 describes how the hyperparameters can be estimated from available training data, using *maximum likelihood estimation* (MLE).

- Section 3.5 highlights how a mrGp with fixed hyperparameters can then be used for prediction. The previously mentioned sections deal with the simplified case of a noise free process.

- Section 3.6, oppositely to the previous sections, addresses a scenario in which the process of interest contains observation error, i.e. a process which presents residual variations, noise or other sources of uncorrelated error.

- Finally, Section 3.7 details a Cholesky factorisation for the mrGp formulation, which allows to improve the numerical stability of the algorithm and which has been implemented on top of the original formulation of the MBA.

3.2 Gaussian process mean

3.2.1 Single response

The prior mean is assumed to belong to a hierarchical structure of linear functions that, given a set of N input data points x, can be written as

$$m(\cdot) = \boldsymbol{H}(\cdot)\boldsymbol{\beta}, \tag{3.1}$$

with regression matrix $\boldsymbol{H}$ and a column vector of regression coefficients $\boldsymbol{\beta}$. The resulting mean column vector m has dimension N. Matrix $\boldsymbol{H} \in \mathbb{R}^{N \times p}$ contains N polynomial functions $f_j(x) : j = 1, \ldots, p$ of degree p. The function takes a set of d input data points, which are denoted as $x = x_1, x_2, \ldots, x_d$ and outputs the value of the regressor polynomial function. Note that the degree of the polynomial function p will depend of how many input data points there are. Different types of functions are implemented as can be seen in Table 3.1

Name	order	p	$f(x)$	
regpoly0	0	1	$f_1(x) = 1$	constant
regpoly1	1	$d + 1$	$f_1(x) = 1, \ldots, f_j(x) = x_{j-1}$	linear
regpoly2	2	$\frac{1}{2}(d + 1)(d + 2)$	$f_q(x)$	quadratic

Table 3.1: Polynomial regression functions for $j = 1, \ldots, p$.

with the above quadratic function $f_q(x)$ defined as

$$\begin{aligned} &f_1(x) = 1 \\ &f_2(x) = x_1, \ldots, f_{d+1}(x) = x_d \\ &f_{d+2}(x) = x_1^2, \ldots, f_{2d+1}(x) = x_1 x_d \\ &f_{2d+2}(x) = x_2^2, \ldots, f_{3d}(x) = x_2 x_d \\ &\ldots \\ &f_p(x) = x_d^2 \end{aligned} \tag{3.2}$$

On the other hand, vector $\boldsymbol{\beta}$ can be defined as $\boldsymbol{\beta}^T = [\beta_1, \ldots, \beta_p]$ with coefficients β_i for each of the $i = 1, 2, \ldots, p$ terms of the linear functions contained in matrix $\boldsymbol{H}$.

3.2.2 Multiple response

For the similar situation where the Gaussian process has to be fitted to q multiple responses at N input data points, the structure of Eq. (3.1) and $\boldsymbol{H}(\cdot)$ remain unchanged. However, the result now is a mean matrix $m \in \mathbb{R}^{N \times q}$, which has one

column for each response that is being approximated. For further details on matrix normal distributions see Carvalho et al. [189]. The regression coefficients column vector $\boldsymbol{\beta}$ also expands into a matrix $\boldsymbol{\beta} \in \mathbb{R}^{p \times q}$, which has the following structure

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_{11} & \dots & \beta_{1q} \\ \vdots & \vdots & \vdots \\ \beta_{p1} & \dots & \beta_{pq} \end{bmatrix}. \tag{3.3}$$

with coefficients β_{ik} for each of the $k = 1, 2, \dots, q$ responses, and for each of the $i = 1, 2, \dots, p$ terms of the linear functions contained in matrix $\boldsymbol{H}$. Thus, the multiple response can be interpreted as a generalisation of the single response case.

Note that sometimes it is necessary to vectorise matrix m into a single column vector to match it with the covariance structure (shown in the next section). The Kronecker product $\otimes$ is the most efficient operator to accomplish the restructuring of the mean vector, and will be used throughout the rest of this work. As an example, Eq. (3.1) can be re-written as

$$m(\cdot) = \boldsymbol{I}_q \otimes \boldsymbol{H}(\cdot)\text{vec}(\boldsymbol{\beta}) \tag{3.4}$$

with identity matrix $\boldsymbol{I}_q$ of dimension q and vec$(\cdot)$ as the vectorising operation. Vectorising $\boldsymbol{\beta}$ will match its dimension with the Kronecker product $\boldsymbol{I}_q \otimes \boldsymbol{H}$. Essentially each mean response is approximated by the same functions $\boldsymbol{H}$ (which are expanded by the Kronecker product) but have different regression coefficients $\boldsymbol{\beta}$, resulting in a Nq column matrix $m = [m^1, m^2, \dots, m^q]^T$.

3.3 Gaussian process covariance

3.3.1 Single response

The prior covariance function for a single response Gaussian process is assumed as the product of a scalar variance σ^2 and a correlation matrix $\boldsymbol{R} \in \mathbb{R}^{N \times N}$ between data points, and can be formulated as

$$\boldsymbol{V}(\cdot, \cdot) = \sigma^2 \boldsymbol{R}(\cdot, \cdot) \tag{3.5}$$

Depending on the nature of the process being approximated different correlation functions can be specified inside each entry of matrix $\boldsymbol{R}$. These are commonly known in the machine learning community as *kernels*. Note also that the correlation matrix of an uncorrelated Gaussian process would be an identity matrix I_N.

Currently the following form is assumed

$$R(\omega, x, x') = \prod_{j=1}^{d} R_j(\omega, x_j - x'_j)$$

where once again a set of d input data points $x = x_1, x_2, \ldots, x_d$, are supplied as an argument of the function, and the output is the required correlation. Similarly to the hierarchy of regression functions there is also a list of correlation functions available

Name	$R_j(\omega, \Delta_j)$	
exp	$\exp(-\omega_j\lvert\Delta_j\rvert)$	
expg	$\exp(-\omega_j\lvert\Delta_j\rvert^{\omega_{n+1}})$,	$0 < \omega_{n+1} \leq 2$
Gauss	$\exp(-\omega_j\Delta_j^2)$	
lin	$\max\{0, 1 - \omega_j\lvert\Delta_j\rvert\}$	
spherical	$1 - 1.5\zeta_j + 0.5\zeta_j^3$,	$\zeta_j = \min\{1, \omega_j\lvert\Delta_j\rvert\}$
cubic	$1 - 3\zeta_j^2 + 2\zeta_j^3$,	$\zeta_j = \min\{1, \omega_j\lvert\Delta_j\rvert\}$
spline	$\varsigma(\zeta_j)$, (3.6)	$\zeta_j = \omega_j\lvert\Delta_j\rvert$

Table 3.2: Correlation functions. $\Delta_j = x_j - x'_j$.

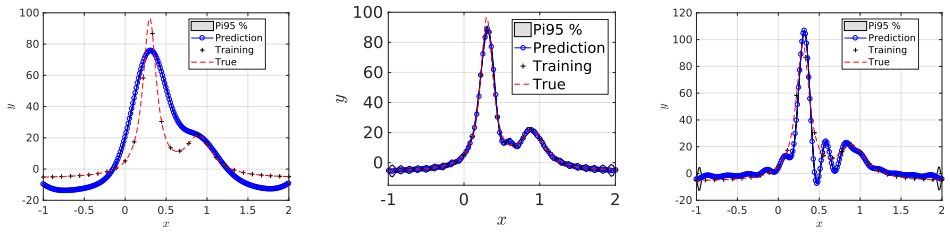
with the above spline model defined as

$$\varsigma(\zeta_j) = \begin{cases} 1 - 15\zeta_j^2 + 30\zeta_j^3 & \text{for} \quad 0 \leq \zeta_j \leq 0.2 \\ 1.25(1 - \zeta_j)^3 & \text{for} \quad 0.2 \leq \zeta_j \leq 1 \\ 0 & \text{for} \quad \zeta_j \leq 1 \end{cases} \tag{3.6}$$

and $\boldsymbol{\omega}_j \;\; j = 1, \ldots, d$ called roughness parameters for each of the d design variables x_i. They represent how roughly the response changes from input point x to point x'. Fig. 3.1 displays the effect of different roughness parameters in the predicted response. Notice how the smoothness of the process decreases as the roughness parameters increase, from a lower value in Fig. 3.1(a) to a value appropriate for the latent function in Fig. 3.1(b) and an higher value in Fig. 3.1(c).

To complement the tabled functions Fig. 3.2 displays most of the presented functions in a graphical format, with distance between points in the abscissa and correlation in the ordinate. As can be seen the functions can either have a parabolic behaviour (GAUSS, CUBIC and SPLINE) near the origin, or linear (EXP, LIN and SPHERICAL).

Also, as can be seen in Fig. 3.3 different correlation functions result in different fitting and prediction intervals. Since the shown hump process is highly nonlinear it is preferable to approximate it with a Gaussian correlation function, whereas a



(a) (b) (c)

Figure 3.1: Approximation of an hump function by a mrGp with different roughness parameters and a Gaussian correlation function: (a) $\omega = 10.00$, (b) $\omega = 31.67$, and (c) $\omega = 50.00$.

linear process would be better approximated with a spherical or linear correlation function.

There is a considerable amount of research dedicated exclusively to the development of efficient correlation functions for Gaussian processes. A comparative study of some of these functions can be seen in Simoen et al. [72], where the spherical and Gaussian models yielded the best results. Typically, engineers use the Gaussian function because of its ease of interpretation and differentiability [146, 190–192]. For the case-studies of this work the Gaussian and linear model generally presented the best results.

As can be seen from the functions of Table 3.2, the correlation of two coincident points $\Delta = 0$ is often one. As shall be seen in the next section, such correlation allows a Gaussian process posterior mean/covariance that yields a perfect interpolation at observed points, and is a convenient structure to approximate computer simulations which are deterministic in nature, always yielding the same output for a fixed input.

3.3.2 Multiple response

Similarly as with the mean function prior for the multiple response case, the prior covariance structure of a mrGp $\boldsymbol{V}$, given a set of N input data points with q responses, is assumed as the product of a variance matrix $\boldsymbol{\Sigma}^2 \in \mathbb{R}^{q\times q}$ and a correlation matrix between data points $\boldsymbol{R} \in \mathbb{R}^{N\times N}$. Once more, the Kronecker product provides the best operator to combine the two terms of the covariance structure, which can be formulated as

$$\boldsymbol{V}(\cdot,\cdot) = \boldsymbol{\Sigma}^2 \otimes \boldsymbol{R}(\cdot,\cdot) \tag{3.7}$$

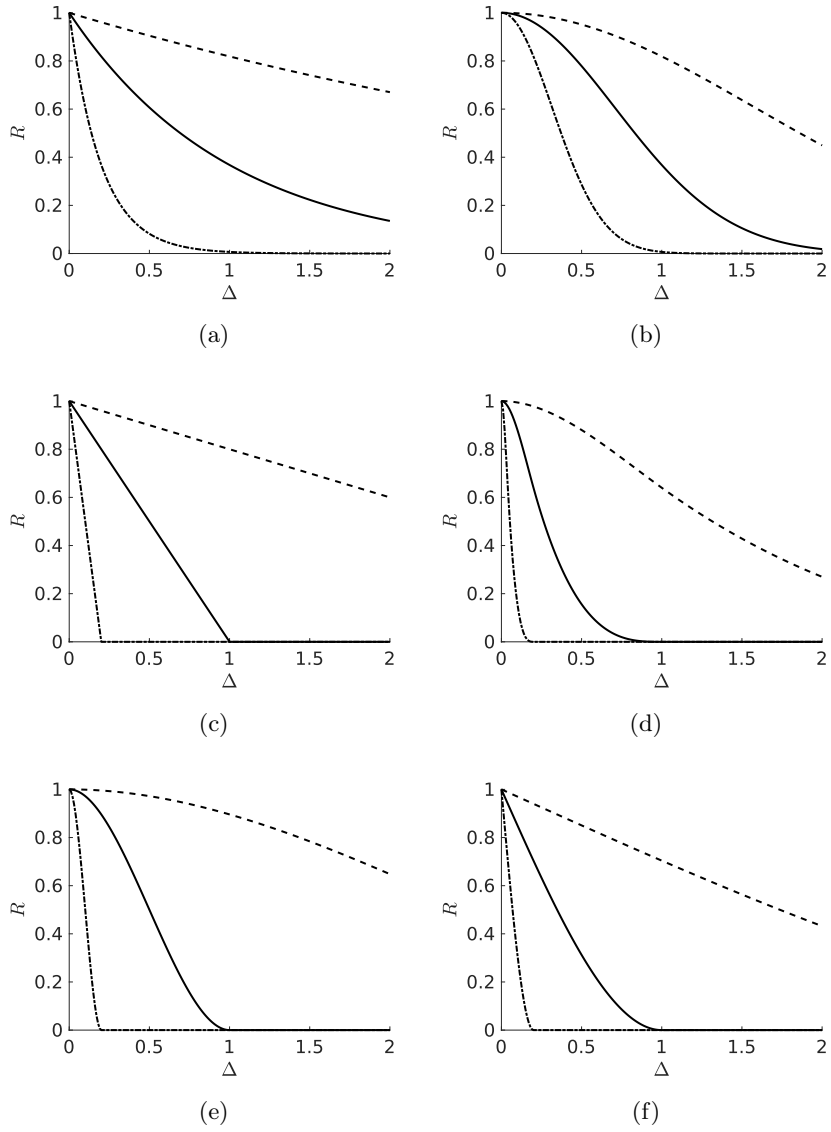
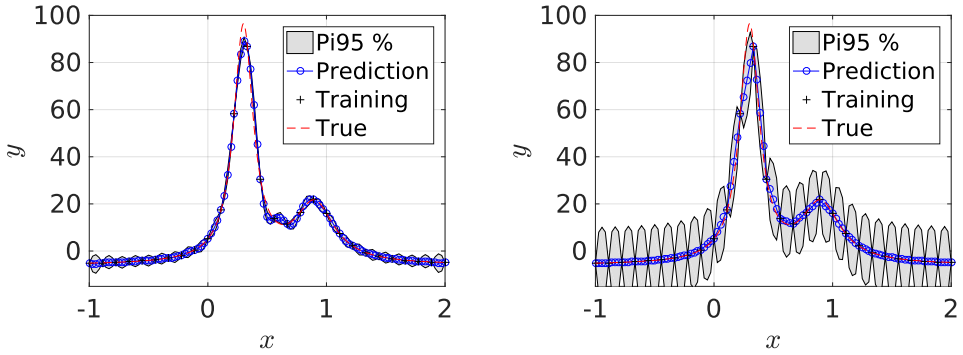


Figure 3.2: Correlation R in the interval $0 \leq \Delta \leq 2$ with roughness parameters $\omega = 0.2, 1, 5$ (dashed, solid and dash-dotted line) for (a) exponential, (b) Gaussian, (c) linear,(d) spline, (e) cubic, and (f) spherical function.

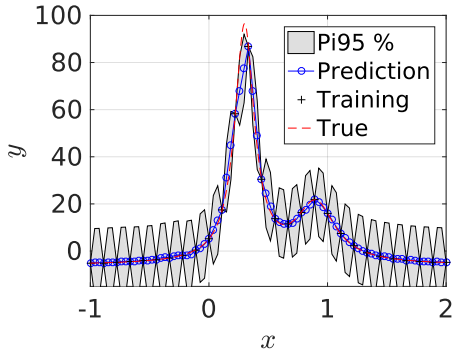
Note the similarity between Eq. (3.7) and Eq. (3.5). Its interpretation is also intuitive, as a separation of a spatial covariance matrix between the q responses that are being fitted, and the temporal correlation between N asynchronous observed points in the input space for stationary processes [185, 193].

However note that the presented covariance assumes that a given input is asynchronous, and that *all* the q output responses have a common roughness pa-



(a)

(b)



(c)

Figure 3.3: Approximation of an hump function by a mrGp with different correlation functions and roughness parameters: (a) Gaussian $\omega = 31.67$, (b) linear $\omega = 0.48$, and (c) spherical $\omega = 1.97$.

rameter $\boldsymbol{\omega}$.

3.4 Estimating hyperparameters

Having established the prior functions of the Gaussian processes for the single and multiple response case it is now possible to address the question of estimating its parameters, here denoted as hyperparameters, from available data. Loosely speaking the Gaussian process is being 'trained' to match supplied data, so that its functions resemble the training data has reliably has possible.

The hyperparameters which have been described in the previous sections $\boldsymbol{\phi} = \{\boldsymbol{\beta}, \boldsymbol{\Sigma}^2, \boldsymbol{\omega}\}$ can be estimated through maximum likelihood estimation (MLE) between the observed data $\boldsymbol{D} = \{\boldsymbol{Y}, \boldsymbol{X}\}$ and the mrGp statistical model. In the subsequent text only the multiple response case will be detailed (from which the

single response case can be readily deduced).

Using MLE is similar to Bayesian estimation, in the sense that a likelihood has also to be built to estimate a hyperparameter of a statistical model. However, the main difference is that MLE does not require us to define our prior beliefs in order to estimate the hyperparameter. Additionally, the hyperparameter attains a fixed deterministic value at the maximum of the likelihood function, whereas Bayesian estimation assumes it as a random variable. MLE is also known as evidence approximation or empirical Bayes [194].

The first step requires labouring a multivariate normal likelihood function of the mrGp, which is analogous to the multivariate normal distribution function found throughout statistical literature

$$f_{\mathbf{X}}(x_1, \ldots, x_N) = \frac{\exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)}{\sqrt{(2\pi)^N|\boldsymbol{\Sigma}|}} \tag{3.8}$$

only differing in the mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, which are replaced by the priors which were presented in the previous sections.

It should be made clear that in the above expression, $\mathbf{x}$ is usually the random variable and $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are fixed statistics, but oppositely, the likelihood assumes that the data is fixed (observed as mentioned before), and that the statistics are variables which values need to be determined.

Note also that a determinant property of the Kronecker product

$$|\mathbf{A} \otimes \mathbf{B}| = |\mathbf{A}|^m |\mathbf{B}|^n . \tag{3.9}$$

has to be used to expand the denominator of Eq. (3.8). The exponent in $|\mathbf{A}|$ is the order of $\mathbf{B}$ and the exponent in $|\mathbf{B}|$ is the order of $\mathbf{A}$.

Hence, the likelihood function of a mrGp is given by

$$\begin{aligned} p(\boldsymbol{Y}|\boldsymbol{\beta}, \boldsymbol{\Sigma}^2, \boldsymbol{\omega}) &= (2\pi)^{-Nq/2}|\boldsymbol{\Sigma}^2|^{-N/2}|\boldsymbol{R}|^{-q/2} \\ &\quad \exp\left\{-\frac{1}{2}\mathrm{vec}(\boldsymbol{Y}-\boldsymbol{H}\boldsymbol{\beta})^T(\boldsymbol{\Sigma}^2 \otimes \boldsymbol{R})^{-1}\mathrm{vec}(\boldsymbol{Y}-\boldsymbol{H}\boldsymbol{\beta})\right\}, \end{aligned} \tag{3.10}$$

and its log-likelihood is

$$\begin{aligned} \ell(\boldsymbol{Y}|\boldsymbol{\beta}, \boldsymbol{\Sigma}^2, \boldsymbol{\omega}) &\propto -\frac{1}{2}\{N\log|\boldsymbol{\Sigma}^2| + q\log|\boldsymbol{R}| \\ &\quad +\mathrm{vec}(\boldsymbol{Y}-\boldsymbol{H}\boldsymbol{\beta})^T(\boldsymbol{\Sigma}^2 \otimes \boldsymbol{R})^{-1}\mathrm{vec}(\boldsymbol{Y}-\boldsymbol{H}\boldsymbol{\beta})\} \end{aligned} \tag{3.11}$$

By differentiating Eq. (3.11) in order to $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}^2$ and equating these deriva-

tives to zero, it is possible to obtain the analytic MLE of these hyperparameters. The following two subsections will detail the derivation and estimation of the regression coefficients $\boldsymbol{\beta}$ and the process variance matrix $\boldsymbol{\Sigma}^2$.

3.4.1 Estimation of regression coefficients

As described above, the likelihood function of a mrGp is multivariate normal, and its mean function is made of regression functions. Fortunately, the solution of the MLE for the particular case is a well known solution (reproduced below).

To determine a weighted least squares solution to the system

$$\boldsymbol{Y} = \boldsymbol{H}\boldsymbol{\beta} + \boldsymbol{\epsilon} \tag{3.12}$$

where $\boldsymbol{\epsilon}$ is a "residual" term with covariance as specified above. Firstly, the generalised measure of the squared distance from $\boldsymbol{Y}$ to $\boldsymbol{H}\boldsymbol{\beta}$ in standard deviation units can be written as

$$(\boldsymbol{Y} - \boldsymbol{H}\boldsymbol{\beta})^T \boldsymbol{R}^{-1} (\boldsymbol{Y} - \boldsymbol{H}\boldsymbol{\beta}), \tag{3.13}$$

which is identical to the univariate normal case $\left(\frac{x-\mu}{\sigma^2}\right) = (x-\mu)(\sigma^2)^{-1}(x-\mu)$. Secondly, by differentiating this equation with respect to $\boldsymbol{\beta}$, and by setting it equal to zero, it can be seen that the minimum of the squares (and its norm) occurs at the $\hat{\boldsymbol{\beta}}$ that satisfies

$$\boldsymbol{H}^T \boldsymbol{R}^{-1} \boldsymbol{H} \hat{\boldsymbol{\beta}} = \boldsymbol{H}^T \boldsymbol{R}^{-1} \boldsymbol{Y}, \tag{3.14}$$

which are also called the *normal equations*. The circumflex above $\boldsymbol{\beta}$ denotes that it is an estimated value. Finally, multiplying the inverse of the first three matrices on both sides the estimate can be isolated as

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{H}^T \boldsymbol{R}^{-1} \boldsymbol{H})^{-1} \boldsymbol{H}^T \boldsymbol{R}^{-1} \boldsymbol{Y}. \tag{3.15}$$

The above shown procedure is analogous to solving $\partial \ell / \partial \boldsymbol{\beta} = 0$, since for the Gaussian case the MLE and the least squares solution is identical. An alternative form of the above estimate can be written as

$$\hat{\boldsymbol{\beta}} = \boldsymbol{W} \boldsymbol{H}^T \boldsymbol{R}^{-1} \boldsymbol{Y}, \tag{3.16}$$

with the Gram-Schmidt matrix inverse, that dictates the numerical stability of the solution defined as

$$\boldsymbol{W} = (\boldsymbol{H}^T \boldsymbol{R}^{-1} \boldsymbol{H})^{-1}. \tag{3.17}$$

Its condition determines the expected accuracy of a solution to the least squares problem. This solution correspond to the classical generalised least squares solution. If the correlation matrix $\boldsymbol{R}$ had been assumed instead as diagonal, i.e. an uncorrelated process, the result would be the ordinary least squares solution.

3.4.2 Estimation of process variance

On the other hand, the MLE of the process variance can be obtained by solving $\partial\ell/\partial\boldsymbol{\Sigma}^2 = 0$, and results in the estimate

$$\hat{\boldsymbol{\Sigma}}^2 = \frac{1}{N-p}(\boldsymbol{Y} - \boldsymbol{H}\hat{\boldsymbol{\beta}})^T \boldsymbol{R}^{-1}(\boldsymbol{Y} - \boldsymbol{H}\hat{\boldsymbol{\beta}}). \tag{3.18}$$

known as generalised sample variance or average of the squared deviations.

3.4.3 Numerical optimisation of the log-likelihood function

By plugging expressions of Eqs. (3.16) and (3.18) into Eq. (3.11) it simplifies into

$$\ell(\boldsymbol{Y}|\boldsymbol{\omega}) \propto -\frac{1}{2}\{N \log|\hat{\boldsymbol{\Sigma}}^2| + q \log|\boldsymbol{R}|\} \tag{3.19}$$

where both terms $\hat{\boldsymbol{\Sigma}}^2$ and $\boldsymbol{R}$ depend of the roughness coefficients $\boldsymbol{\omega}$. This function has to be maximised numerically in order to estimate the values of $\boldsymbol{\omega}$.

3.5 Prediction mrGp

Finally, it is possible to generate predictions of the latent function $\boldsymbol{y}$, using a Gaussian process which has been trained with observed data $\boldsymbol{D} = \{\boldsymbol{Y}, \boldsymbol{X}\}$ and from which hyperparameters $\boldsymbol{\phi}$ have been estimated. To this end, let us assume that s unobserved point(s) $\boldsymbol{x}_* \in \mathbb{R}^{s\times d}$, are also provided. The objective is to compute the respective predictions $\boldsymbol{y}(\boldsymbol{x}_*)$ of one or q responses, for the single or the multiple response case, respectively. Thus, $\boldsymbol{y}$ can either be a s column vector or a $s \times q$ matrix.

3.5.1 Single response

Since $\boldsymbol{y}(\boldsymbol{x}_*)$ is assumed to belong to the same distribution as the dataset $\boldsymbol{D}$, i.e. the same latent function, its prior PDF is

$$\boldsymbol{y}(\boldsymbol{x}_*)|\boldsymbol{\beta}, \sigma^2, \boldsymbol{\omega} \sim \mathcal{N}(\boldsymbol{h}(\boldsymbol{x}_*)\boldsymbol{\beta}, \sigma^2 R(\boldsymbol{x}_*, \boldsymbol{x}'_*)) \tag{3.20}$$

with the definitions of $R(\cdot,\cdot), \sigma^2$ and $\boldsymbol{h}(\cdot)\boldsymbol{\beta}$ detailed above. Similarly, the observed data $\boldsymbol{D}$ is assumed to follow a multivariate normal distribution as

$$\boldsymbol{D}|\boldsymbol{\beta}, \sigma^2, \boldsymbol{\omega} \sim \mathcal{N}(\boldsymbol{H}(\boldsymbol{X})\boldsymbol{\beta}, \sigma^2 \boldsymbol{R}(\boldsymbol{X}, \boldsymbol{X}')). \tag{3.21}$$

Note that matrices R and $\boldsymbol{h}$ are denoted differently from matrices $\boldsymbol{R}$ and $\boldsymbol{H}$, just so that it is clear to which distribution each matrix belongs to.

The situation of interest now is to obtain the conditional posterior distribution of $\boldsymbol{y}(\boldsymbol{x}_*)$ based on the dataset $\boldsymbol{D}$. Since both distributions are multivariate normal, the conditional case is available from standard normal theory (see Appendix A for a full proof) as

$$\boldsymbol{y}(\boldsymbol{x}_*)|\boldsymbol{\beta}, \sigma^2, \boldsymbol{\omega}, \boldsymbol{D} \sim \mathcal{N}(\boldsymbol{m}_*, \boldsymbol{V}_*) \tag{3.22}$$

where the mean vector of this distribution is

$$\boldsymbol{m}_* = \boldsymbol{h}\boldsymbol{\beta} + \mathcal{R}^T \boldsymbol{R}^{-1}(\boldsymbol{Y} - \boldsymbol{H}\boldsymbol{\beta}) \tag{3.23}$$

and its covariance matrix is

$$\boldsymbol{V}_* = \sigma^2\{R - \mathcal{R}^T \boldsymbol{R}^{-1}\mathcal{R}\} \tag{3.24}$$

with $\mathcal{R} = [R(\boldsymbol{x}_*, \boldsymbol{x}_1), \ldots, R(\boldsymbol{x}_*, \boldsymbol{x}_N)]^T \in \mathbb{R}^{N\times s}$ defined as a relational correlation matrix, which maps the correlation between the dataset points and the unobserved points. It is worth noting, that the first term in Eq. (3.23) $\boldsymbol{h}\boldsymbol{\beta}$ is a mean of the predicted values, and the last arguments $\boldsymbol{R}^{-1}(\boldsymbol{Y} - \boldsymbol{H}\boldsymbol{\beta})$ are fixed, because they depend only of the training dataset/hyperparameters. When multiplied by matrix $\mathcal{R}$ the last term acts as a set of weights, which adjust the Gaussian process according to training data. An example of predictions made by a Gaussian process with and without training data is presented in Section 4.4, to illustrate the relevance of each of the terms in the above equations.

Notice however, that the multivariate normal distribution in Eq. (3.22) depends on the hyperparameters $\boldsymbol{\phi} = \{\boldsymbol{\beta}, \sigma^2, \boldsymbol{\omega}\}$. After estimating those hyperparameters as described in Section 3.4, it is necessary to integrate Eq. (3.22) with respect to the MLE of $\boldsymbol{\beta}$ and σ^2 to obtain the distribution of $\boldsymbol{y}$ conditional on the roughness parameters $\boldsymbol{\omega}$ and data $\boldsymbol{D}$ only. This distribution that closes the description of the Gaussian process prediction model, is given by a t-student distribution defined by

$$\boldsymbol{y}(\boldsymbol{x}_*)|\boldsymbol{\omega}, \boldsymbol{D} \sim \mathcal{T}(\boldsymbol{m}^{**}, \hat{\sigma}^2 \boldsymbol{R}^{**}) \tag{3.25}$$

with

$$\boldsymbol{m}^{**} = \boldsymbol{h}\hat{\boldsymbol{\beta}} + \mathcal{R}^T\boldsymbol{R}^{-1}(\boldsymbol{Y} - \boldsymbol{H}\hat{\boldsymbol{\beta}}) \tag{3.26}$$

$$\boldsymbol{R}^{**} = R - \mathcal{R}^T\boldsymbol{R}^{-1}\mathcal{R} + [\boldsymbol{h}^T - \boldsymbol{H}^T\boldsymbol{R}^{-1}\mathcal{R}]^T[\boldsymbol{H}^T\boldsymbol{R}^{-1}\boldsymbol{H}]^{-1}[\boldsymbol{h}^T - \boldsymbol{H}^T\boldsymbol{R}^{-1}\mathcal{R}], \tag{3.27}$$

and roughness parameters $\boldsymbol{\omega}$ estimated by maximisation of Eq. (3.19) and plugged into Eq. (3.25) to achieve the prediction at the new input data points.

As an example, in Fig 3.4 an increasing number of training data sets are used to fit a Gaussian process against a latent function (dashed line). Note that the posterior function predictions are coincident with training data $\boldsymbol{y} = \boldsymbol{y}_1, \ldots, \boldsymbol{y}_q$ at $\boldsymbol{x}_* = \boldsymbol{x}_1, \ldots, \boldsymbol{x}_N$ and the predicted points covariance goes to zero, i.e. perfect interpolation applies.

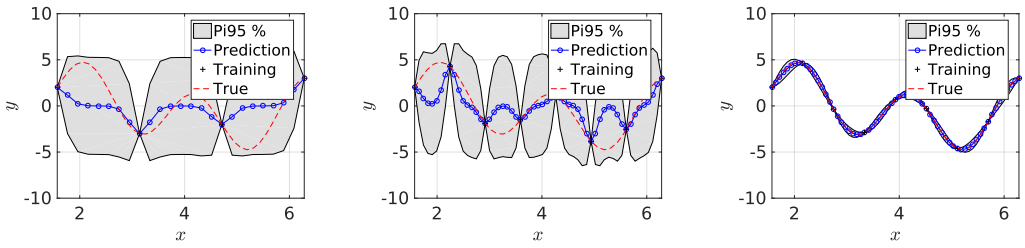


Figure 3.4: As the number of training data points (plus dots) increases (left to right) a mean function (solid line) interpolates perfectly these points and the uncertainty cloud (grey area) shrinks further around it, converging to the true process function (dashed line).

3.5.2 Multiple responses

In a similar way as to what has been shown in previous sections, the posterior mrGp formulation is a natural expansion of the single response case.

The conditional posterior distribution of $\boldsymbol{y}(\boldsymbol{x}_*)$ based on the dataset $\boldsymbol{D}$ can be defined as

$$\boldsymbol{y}(\boldsymbol{x}_*)|\boldsymbol{\beta}, \boldsymbol{\Sigma}^2, \boldsymbol{\omega}, \boldsymbol{D} \sim \mathcal{N}, \boldsymbol{V}_*) \tag{3.28}$$

where the mean vector is

$$\boldsymbol{m}_* = \boldsymbol{h}\boldsymbol{\beta} + \mathcal{R}^T\boldsymbol{R}^{-1}(\boldsymbol{Y} - \boldsymbol{H}\boldsymbol{\beta}) \tag{3.29}$$

and the covariance matrix is

$$\boldsymbol{V}_* = \boldsymbol{\Sigma}^2 \otimes \{R - \mathcal{R}^T \boldsymbol{R}^{-1} \mathcal{R}\} \tag{3.30}$$

Analogous to what has been shown for the single response case, the aforementioned distribution can be marginalised in order to the hyperparameters, resulting in a t-student distribution defined by

$$\boldsymbol{y}(\boldsymbol{x}_*)|\boldsymbol{\omega}, \boldsymbol{D} \sim \mathcal{T}(\boldsymbol{m}^{**}, \hat{\boldsymbol{\Sigma}}^2 \otimes \boldsymbol{R}^{**}) \tag{3.31}$$

with

$$\begin{aligned} \boldsymbol{m}^{**} &= \boldsymbol{h}\hat{\boldsymbol{\beta}} + \mathcal{R}^T \boldsymbol{R}^{-1}(\boldsymbol{Y} - \boldsymbol{H}\hat{\boldsymbol{\beta}}) && (3.32) \\ \boldsymbol{R}^{**} &= R - \mathcal{R}^T \boldsymbol{R}^{-1} \mathcal{R} \\ &+ [\boldsymbol{h}^T - \boldsymbol{H}^T \boldsymbol{R}^{-1} \mathcal{R}]^T [\boldsymbol{H}^T \boldsymbol{R}^{-1} \boldsymbol{H}]^{-1} [\boldsymbol{h}^T - \boldsymbol{H}^T \boldsymbol{R}^{-1} \mathcal{R}], && (3.33) \end{aligned}$$

and roughness parameters $\boldsymbol{\omega}$ estimated by maximisation of Eq. (3.19) and plugged into Eq. (3.31) to achieve the prediction at the new input data points.

In the next section a variant of this formulation, that allows to consider observation errors is detailed. This is important when the mrGp is being fitted against measured data, which presents residual variations, noise or other sources of uncertainty when at the same conditions of observation.

3.6 Observation error

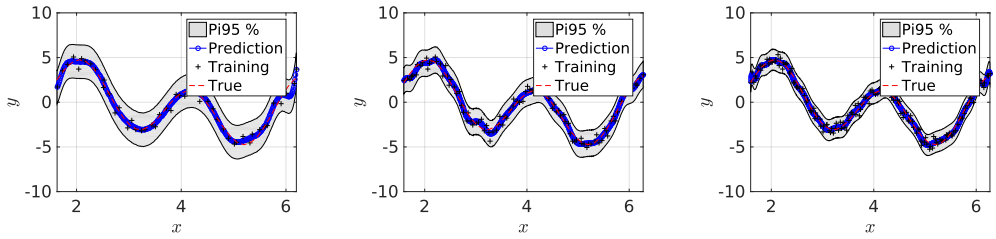
In the previous sections, a mrGp formulation that allows to fit a deterministic process, with available data was detailed. Now the situation of interest is a process with observation error, i.e. a process which presents uncertainty at the same conditions of observation.

An example of such process is shown in Fig. 3.5, where a signal with a signal-to-noise ratio SNR of 35 dB is fitted by a Gaussian process. The SNR is defined as

$$\mathrm{SNR}_{\mathrm{dB}} = 10 \log_{10} \left(\frac{\sigma^2_{\mathrm{signal}}}{\sigma^2_{\mathrm{noise}}} \right), \tag{3.34}$$

where σ_{signal} and σ_{noise} represent the variance of the signal of interest and noise, respectively.

The observation error is assumed as zero-mean Gaussian. Hence the mean structure of Section 3.2 remains unchanged. On the other hand an additional term



(a) (b) (c)

Figure 3.5: Gaussian process approximation when training data consists of noisy observations, SNR of 35 dB, for (a) 50, (b) 100 and (c) 150 dataset points.

was added to the covariance structure of Eq. (3.7) as

$$\boldsymbol{V}(\cdot,\cdot) = \boldsymbol{\Sigma}^2 \otimes \boldsymbol{R}(\cdot,\cdot) + \boldsymbol{\Lambda} \otimes \boldsymbol{I}_n, \tag{3.35}$$

where $\boldsymbol{I}_N$ is the identity matrix of dimension N and $\boldsymbol{\Lambda} \in \mathbb{R}^{q\times q}$ is a variance matrix of the observation error for each response. Implicitly, Eq. (3.35) separates the behaviour of the process in a correlated and uncorrelated structure, between the error at different input data points.

To estimate the hyperparameters $\boldsymbol{\phi} = \{\boldsymbol{\beta}, \boldsymbol{\Sigma}^2, \boldsymbol{\Lambda}, \boldsymbol{\omega}\}$ with MLE, only the regression coefficients $\boldsymbol{\beta}$ do have an analytic expression, given by

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{I}_q \otimes \boldsymbol{H}^T \boldsymbol{V}^{-1} \boldsymbol{I}_q \otimes \boldsymbol{H})^{-1} \boldsymbol{I}_q \otimes \boldsymbol{H}^T - \boldsymbol{V}^{-1}\mathrm{vec}(\boldsymbol{Y}) \tag{3.36}$$

while the remaining hyperparameters have to be estimated by numerical optimisation methods. The log-likelihood function that has to be maximised in order to do so is given by

$$\ell(\boldsymbol{Y}|\boldsymbol{\phi}) \propto -\frac{1}{2}\{\log|\boldsymbol{V}| + \mathrm{vec}(\boldsymbol{Y} - \boldsymbol{H}\hat{\boldsymbol{\beta}})^T \boldsymbol{V}^{-1}\mathrm{vec}(\boldsymbol{Y} - \boldsymbol{H}\hat{\boldsymbol{\beta}})\}. \tag{3.37}$$

Finally, and similarly to Section 3.5, the prediction at an unobserved point $\boldsymbol{x}_*$ is given by a t-student distribution defined by

$$\boldsymbol{y}(\boldsymbol{x}_*)|\boldsymbol{\omega}, \boldsymbol{D} \sim \mathcal{T}(\boldsymbol{m}^{**}, \boldsymbol{V}^{**}) \tag{3.38}$$

where

$$\begin{aligned} \boldsymbol{m}^{**} &= \boldsymbol{h}\hat{\boldsymbol{\beta}} + \mathcal{V}^T\boldsymbol{V}^{-1}\text{vec}(\boldsymbol{Y} - \boldsymbol{H}\hat{\boldsymbol{\beta}}) \qquad (3.39) \\ \boldsymbol{V}^{**} &= \hat{\boldsymbol{\Sigma}}^2 \otimes R + \hat{\boldsymbol{\Lambda}} \otimes \boldsymbol{I}_s - \mathcal{V}^T\boldsymbol{V}^{-1}\mathcal{V} \\ &+ [\boldsymbol{h}^T - \boldsymbol{H}^T\boldsymbol{V}^{-1}\mathcal{V}]^T[\boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H}]^{-1}[\boldsymbol{h}^T - \boldsymbol{H}^T\boldsymbol{V}^{-1}\mathcal{V}] \qquad (3.40) \end{aligned}$$

with $\mathcal{V} = \{\hat{\boldsymbol{\Sigma}}^2 \otimes [R(\boldsymbol{x}_*, \boldsymbol{x}_1), \ldots, R(\boldsymbol{x}_*, \boldsymbol{x}_N)]\}^T \in \mathbb{R}^{Nq \times s}$ as a relational covariance matrix. Note that the regression matrix $\boldsymbol{H}$ in Eq. (3.40) should have been expanded with the Kronecker product, as shown in Eq. (3.4) but for the sake of clarity it has been omitted.

If a single input site is required, the above formulas compact directly, including $\mathcal{V}$ which becomes a $\mathbb{R}^{Nq \times 1}$ vector. Finally, the author would like to comment the subtle differences between Eq. (3.40) and Eq. (3.33). It can be noted that the additional term $\hat{\boldsymbol{\Lambda}} \otimes \boldsymbol{I}_s$ in Eq. (3.40) is the additional variance due to observation error; while all the other terms are equivalent on both equations (note that the correlation matrices in Eq. (3.33) are multiplied by the process variance of Eq. (3.31)).

3.7 Cholesky factor decomposition

As shown in the previous sections, Gaussian process can be used for efficient regression of stochastic processes, with or without observation error. The present section highlights a mathematical artifact which can be used to improve the numerical stability of the mrGp fitting process. Although Rasmussen and Williams [183] have already detailed such operation, it has been restricted to the single response case, while the author has generalised it for the approximation of multiple responses.

Since the latent function that is approximated by a mrGp is assumed as multivariate normal, its likelihood function has a well known form, which can be expressed as

$$\begin{aligned} \ln p(\boldsymbol{Y}|\boldsymbol{\phi}) &= -\frac{N}{2}\ln(2\pi) - \frac{N}{2}\ln|\boldsymbol{V}| \qquad (3.41) \\ &\quad -\frac{1}{2}\text{vec}(\boldsymbol{Y} - \boldsymbol{H}\boldsymbol{\beta})^T\boldsymbol{V}^{-1}\text{vec}(\boldsymbol{Y} - \boldsymbol{H}\boldsymbol{\beta}), \end{aligned}$$

where $\boldsymbol{Y}$ and N are the data to be approximated and its dimension, $\boldsymbol{\phi}$ are the hyperparameters, vec($\bullet$) is the vectorisation (stacking of the columns) and $\boldsymbol{V}$ and $\boldsymbol{H}\boldsymbol{\beta}$ are a covariance and mean function of the mrGp, respectively. For more details of these terms the reader is referred to Sections 3.2 and 3.3.

However the form of Eq. (3.41) is not practical from a numerical point of

view, since it requires a covariance matrix inversion in the last term. The dimension of matrix $\boldsymbol{V}$ can become very large creating numerical stability problems. However, since $\boldsymbol{V}$ is symmetric and positive-definite, it is implicit that it can be written in factorised form [195] as

$$\boldsymbol{V} = \boldsymbol{C}\boldsymbol{C}^T, \tag{3.42}$$

with $\boldsymbol{C}$ a lower triangular matrix. If we then consider that the regression problem which is being solved, translating similarity or equality between a hierarchy of regression functions and available data can be transformed as follows

$$\widetilde{\boldsymbol{H}}\boldsymbol{\beta} \simeq \widetilde{\boldsymbol{Y}} \tag{3.43}$$

with

$$\boldsymbol{C}\widetilde{\boldsymbol{H}} = \boldsymbol{H}, \qquad \boldsymbol{C}\widetilde{\boldsymbol{Y}} = \boldsymbol{Y}, \tag{3.44}$$

Eq. (3.41) then simplifies into

$$\ln p(\boldsymbol{Y}|\boldsymbol{\phi}) = -\frac{N}{2}\ln(2\pi) - N\ln|\boldsymbol{C}| - \frac{1}{2}\mathrm{vec}(\widetilde{\boldsymbol{Y}} - \widetilde{\boldsymbol{H}}\boldsymbol{\beta})^T\mathrm{vec}(\widetilde{\boldsymbol{Y}} - \widetilde{\boldsymbol{H}}\boldsymbol{\beta}), \tag{3.45}$$

which is considerably more stable and efficient than the original formula.

As shown in the previous sections, the used covariance matrix is expressed with a separable form $\boldsymbol{V} = \boldsymbol{\Sigma}^2 \otimes \boldsymbol{R}$, where $\otimes$ is the Kronecker product, $\boldsymbol{\Sigma}^2$ is a covariance between measured responses, and $\boldsymbol{R}$ is a correlation matrix between data points. This form does not however affect any of the above-mentioned simplifications. For more details on properties of the Cholesky factorisation and the Kronecker product the reader is referred to Schäcke [196].

In Chapter 4 the MBA will be presented in detail and several log-likelihood functions, such as Eq. (3.41), will be detailed for each module of the formulation. More stable functions as in Eq. (3.45) have been deduced and implemented for each of the MBA modules and are shown in Appendix C. Such implementation is a core novelty of the present thesis and is believed to improve the applicability of the MBA to quantification of uncertainties in SHM.

Chapter 4

Modular Bayesian uncertainty quantification

4.1 Introduction

In this chapter, the core methodology proposed for uncertainty quantification in SHM, the MBA, is described in detail. Additionally to the factorisation of its mrGps, which has been described in Section 3.7, the original formulation has been expanded to consider multiple calibration parameters. Such considerations imply a complete re-implementation of the original algorithm. It should be noted for example, that Arendt et al. [146] were unable to use their implementation of the MBA for calibration of more than two responses, and as shall be shown in the next chapters, examples with eight responses have been solved with the MBA. Moreover, Arendt et al experienced computational instabilities when training a mrGp with only 16 simulated data points, and in the examples shown in the current chapter 36 simulated data points and above have been used. It is plausible to assume that such improvements are due to the factorisation detailed in Section 3.7. The consequent advantages will become more evident in Chapter 5 and 6. Finally, the differences relatively to other methodologies, such as the classical and hierarchical Bayesian frameworks, will also be highlighted when appropriate.

As described in Section 2.5, the MBA resorts heavily on mrGps for uncertainty quantification. Hence, it is recommended for the reader to be acquainted with mrGps and the contents of Chapter 3 before moving forward. The diagram in Fig. 4.1 highlights two scenarios of data interpretation for SHM, where simulated data $\boldsymbol{X}^m, \boldsymbol{Y}^m, \boldsymbol{\theta}^m$ and experimental data $\boldsymbol{X}^e, \boldsymbol{Y}^e$ are to be used for MSD and st-id. In both cases, the process is broken into a modular structure which processes the

input data, according with the underlying objective/assumptions, and estimates a sensor configuration or calibrates a physics-based model. It should be noted that the MSD modules are derived semi-directly from the st-id ones, and therefore, only the modules which comprise the st-id problem will be detailed in this chapter. Finally, note that sensor configurations obtained by the MSD framework, although tailored for st-id, do not necessarily require a st-id performed by the MBA.

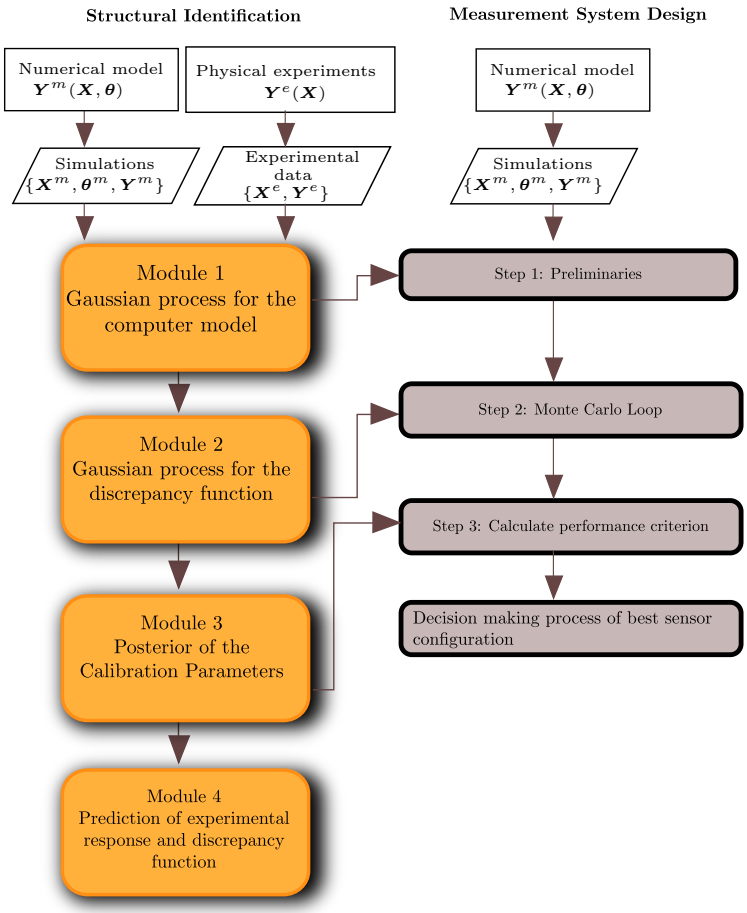


Figure 4.1: Diagram of the relationship between the proposed methodology and a structural identification and measurement system design algorithm. Adapted from [147].

The current chapter is organised as follows: Section 4.2 formalises Bayes' theorem and the involved elements of Bayesian inference. Section 4.3 describes the main equation used to describe the data and quantify the existent uncertainties; Section 4.4 reviews some core aspects of Gaussian processes and how they are used to aid in the uncertainty quantification process of the methodology; Section 4.5 details the approach used for solving the highlighted formulation and highlights novel

developments relatively to the original formulation; and finally Section 4.7 presents the application of the methodology on several simulated examples of a cantilever beam, illustrating the algorithm performance under the presence of: noise, multiple calibration parameters and multiple design-variables.

4.2 Bayes' theorem

A core assumption when using Bayesian inference is that the parameters to be estimated are random variables, with an associated probability density function (PDF). Assuming that a model that describes the data $\boldsymbol{D}$ depends upon parameters $\boldsymbol{\theta}$ that have to be identified, Bayes' Theorem states that

$$\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{marginal likelihood}} \qquad p(\boldsymbol{\theta}|\boldsymbol{D}) = \frac{p(\boldsymbol{D}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{\int p(\boldsymbol{D}|\boldsymbol{\theta})p(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta}} \tag{4.1}$$

where $p(\boldsymbol{\theta}|\boldsymbol{D})$ is the posterior distribution of $\boldsymbol{\theta}$, $p(\boldsymbol{\theta})$ its prior and $p(\boldsymbol{D}|\boldsymbol{\theta})$ is the likelihood function. The denominator is called the marginal likelihood and its purpose is to scale the posterior PDF integration to one, i.e. make it a proper PDF.

Throughout this work the prior distribution will be assumed either as a uniform non-informative PDF, or as a Gaussian PDF, and the likelihood is based on a comprehensive probabilistic data model, which will be presented in Section 4.3. Although in several examples throughout this thesis non-informative priors are used, they do not adhere to the principle of indifference, and therefore, the MBA is genuinely Bayesian and not least squares estimation (see Section 2.2.6 for more details). Note that whenever plots of these three PDFs are shown the likelihood will not be proper. As an example, if the Young's modulus of the beam represented in Fig. 2.5 is identified with the MBA its likelihood with a proper PDF is represented in Fig. 4.2(a), oppositely to a non-proper PDF in Fig. 4.2(b).

4.3 Data model

4.3.1 Observation equation

A given continuous process $\boldsymbol{\xi}$ has n observations of q responses $\boldsymbol{Y}^e \in \mathbb{R}^{n\times q}$ (e stands for experiment) and is dependent on d design variables $\boldsymbol{X}^e \in \mathbb{R}^{n\times d}$. The observation equation can be written as follows

$$\boldsymbol{Y}^e(\boldsymbol{X}^e) = \boldsymbol{\xi}(\boldsymbol{X}^e) + \boldsymbol{\varepsilon} \tag{4.2}$$

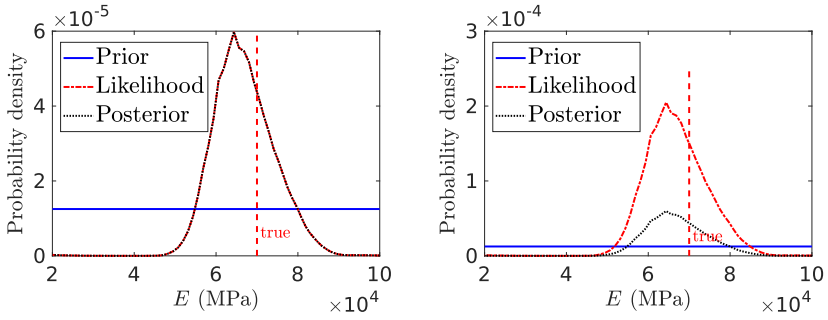


Figure 4.2: Proper (a) vs non-proper (b) PDF. Note that in these plots the posterior is always proper, i.e. its integration equates to one.

where $\boldsymbol{\varepsilon}^T = [\varepsilon_1, \ldots, \varepsilon_n]$ is an observation error that is assumed to follow a zero-mean uncorrelated stationary Gaussian distribution $\mathcal{N}(\boldsymbol{O}, \boldsymbol{\Lambda})$. By its definition, the observation error covariance matrix $\boldsymbol{\Lambda}$ is strictly diagonal.

4.3.2 Physics-based model

On the other hand, the unobservable process $\boldsymbol{\xi}(\boldsymbol{X}^e)$ has to be described using a physics-based model $\boldsymbol{Y}^m$ (termed model for the sake of brevity) as follows

$$\boldsymbol{\xi}(\boldsymbol{X}^e) = \boldsymbol{Y}^m(\boldsymbol{X}^e, \boldsymbol{\theta}) + \boldsymbol{\delta}(\boldsymbol{X}^e), \tag{4.3}$$

where $\boldsymbol{Y}^m(\boldsymbol{X}^e, \boldsymbol{\theta})$ is the model output, dependent of an r-dimensional vector of calibration parameters $\boldsymbol{\theta}$, and $\boldsymbol{\delta}(\boldsymbol{X}^e)$ is a discrepancy function that translates the difference between the true process and the model. The model is assumed as deterministic, i.e. for the same input it will always yield the same output. Note that unlike the classical or hierarchical Bayesian frameworks the discrepancy function *is not* assumed as a zero mean uncorrelated Gaussian. As mentioned previously, the calibration parameters $\boldsymbol{\theta}$ are generic, i.e. they need to be considered independently of the type of analysis which is carried out by the physics-based model and do affect its structural output, e.g. Young's modulus, stiffness of a soil's foundation, etc. On the other hand, the design variables $\boldsymbol{X}$ assume relevance in a specific situation under analysis, by representing e.g. temperature fluctuations in a thermal analysis, humidity in a hygrothermal analysis, wind in a fluid dynamic simulation, time in a dynamic analysis, traffic, etc. Note that the external effects represented by the design variables are not inherent to the structural system. They depend solely of the particular analysis that the structural system is subjected to.

Now replacing (4.3) in (4.2) results in

$$\boldsymbol{Y}^e(\boldsymbol{X}^e) = \boldsymbol{Y}^m(\boldsymbol{X}^e, \boldsymbol{\theta}^*) + \boldsymbol{\delta}(\boldsymbol{X}^e) + \boldsymbol{\varepsilon} \tag{4.4}$$

which is the main equation of model calibration. Eq. (4.4) is analogous to the formulations of the classical and hierarchical Bayesian frameworks, although it includes the design variables $\boldsymbol{X}^e$ which allow to consider other external effects. There are an infinite number of solutions of Eq. (4.3). For different values of $\boldsymbol{\theta}$ there will always be a discrepancy function that matches that particular model instance with $\boldsymbol{\xi}(\boldsymbol{X}^e)$. However the main interest is an ideal state of the model, where $\boldsymbol{\theta}$ attain their true physical values, and the discrepancy function highlights the actual deficiencies of the model. These unknown parameter values $\boldsymbol{\theta}^*$ are designated as *true parameters*.

A scenario where all the quantities $\boldsymbol{\theta}^*$, $\boldsymbol{\delta}$ and $\boldsymbol{\varepsilon}$ are illustrated can be seen in Fig. 4.3. The plot of Fig. 4.3(a) represents the vertical velocity of a viaduct's mid-span during passage of a highspeed train. However the data acquisition system used to obtain such data has a limited precision, and is sensitive to interference from nearby transmission towers, thus presenting some observation error, as can be seen in Fig. 4.3(b). Finally, an SHM analyst has developed a FE model of the viaduct dynamics, with the goal of interpreting all the monitored data and develop future predictions of its structural behaviour. Unfortunately the analyst is unable to calibrate his FE model, mainly because of some unknown parameters $\boldsymbol{\theta}$ which represent the stiffness of soil which surrounded the viaduct and damping of a layer of ballast in the railbed. If the true values of these parameters $\boldsymbol{\theta}^*$ were plugged into the FE model a simulation as shown in Fig. 4.3(c) would be obtained. Notwithstanding such result, when comparing the simulations with the monitored data some discrepancies would still be observable, as shown in Fig. 4.3(d). Finally, it should be mentioned that such discrepancies would be very useful in order to update the FE model. For example, the discrepancy plot shows some velocity frequency ripples which grow up to 4.28 s, decreasing thereafter, which indicates e.g. an interaction effect between the rail and the train wheel, some nonlinear effect of the soil or ballast which triggers high frequency modes of the structural output, etc.

Having formulated a general comprehensive formulation for model calibration, it is now necessary to labour a probabilistic framework which allows to effectively quantify its uncertainties. The next section highlights some relevant features of mrGps used to quantify the uncertainties of Eq. (4.4). Several references are made to Chapter 3 where appropriate. Finally, in Section 4.5 the modular approach which solves the equation is presented.

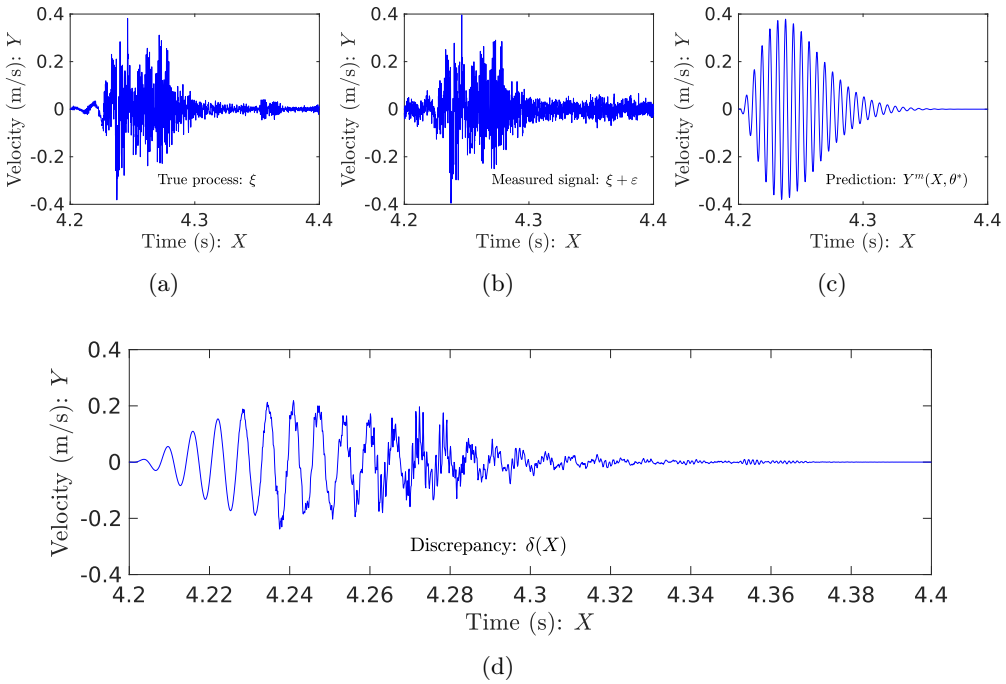


Figure 4.3: Arguments of data model Eq. (4.4), with a true process (a), monitored data (b), predicted data (c) and discrepancy function (d).

4.4 Uncertainty quantification with multiple response Gaussian processes

A short summary of Gaussian processes shall be detailed in this section, along with a discussion of their suitability to approximate the model and discrepancy function detailed in Section 4.3.2.

As detailed in Chapter 3 a Gaussian process is a nonlinear statistical model, which can be used to approximate a Gaussian multivariate normal distribution over a set of functions, based on available training data $\boldsymbol{D} = \{\boldsymbol{X}, \boldsymbol{Y}\}$ and an assumed mean and covariance structures. Chapter 3 also highlighted the extension of a single response Gaussian process into a mrGp case. An example is illustrated in Fig. 4.4, where a set of training data was used to condition a multivariate normal distribution over a set of functions (sinusoidal and hump). Note that the first plots (a) and (d) are prior functions with a constant mean $\boldsymbol{H\beta}$, variance Σ^2 and no training data. The approximation process is also known as regression or fitting, and consists in estimating parameters of the statistical model (hyperparameters) that provide the best fit with the data. In literature, the nomenclature hyperparameters is usually

associated with parameters of probabilistic models, e.g. parameters of a likelihood function, as opposed to input model parameters of an FE model.

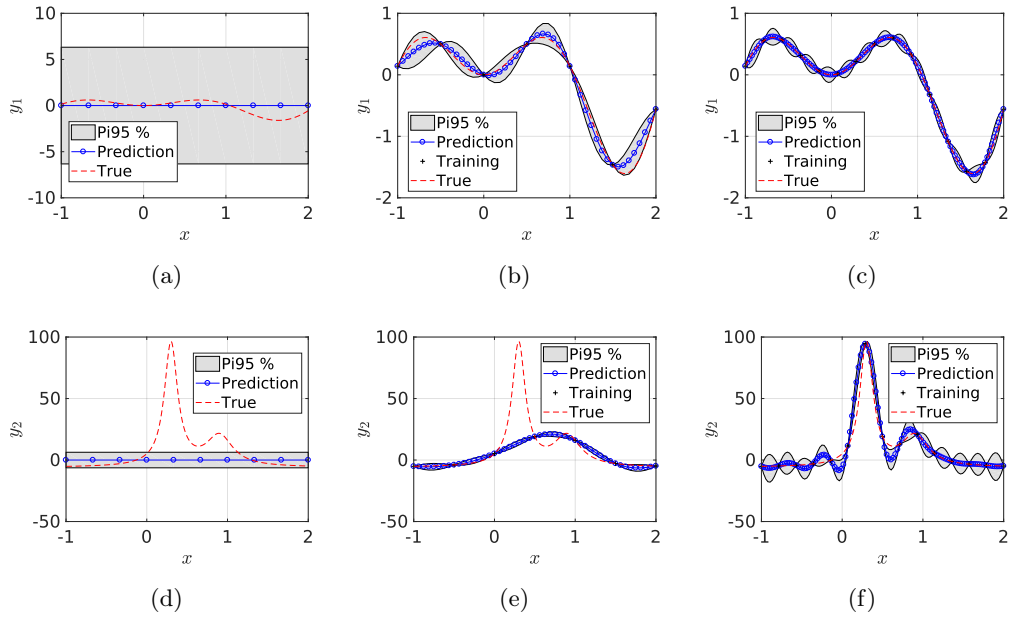


Figure 4.4: Example of mrGp regression of a sinusoidal (a), (b) and (c) and a hump (d), (e) and (f) function. As the number of training data points (plus dots) increases (left to right) a mean function (solid line) interpolates perfectly these points and the uncertainty cloud (grey area) shrinks further around it, converging to the true process function (dashed line).

A Gaussian process also conceptualises Bayesian probabilities, in the sense that the prior mean and covariance functions are being conditioned by the observed data, e.g. see Figs. 4.4(d) and 4.4(e). The resulting posterior distribution can be formulated according to specific requirements. For example, it is possible to consider a perfect interpolation of the available data (relevant when approximating simulations), or alternatively when dealing with noisy/residual measurements, an additional term can be added to the covariance structure to properly account for the additional observation error (see Section 3.3 and 3.6 for further details).

Finally, if a computationally expensive model is being approximated, new realisations of this model can be generated at unobserved regions with associated confidence intervals and a reduced computational effort, e.g., as can be seen by the predictions in between training points in Fig. 4.4(b). When sampling a model parameter space, e.g. in a Latin-hypercube configuration, the number of required samples increases exponentially with the dimensionality of the problem, justifying a faster surrogate model.

All of these reasons justify the use of Gaussian processes as a reliable framework for uncertainty quantification, and as a surrogate of the model and discrepancy function response surfaces. Since in this work several responses $\boldsymbol{Y}(\boldsymbol{X})$ are being considered at the same time, mrGps with the improvements discussed in Section 3.7 will be used primarily.

Reiterating, there is a process with q observed responses at n discrete points, dependent on d design variables. This process has to be described using: a model that takes as input the design variables and r unknown structural parameters; and a discrepancy function (also unknown). Furthermore the observed responses exhibit a measurement error which has also to be estimated. To aid in this quantification, mrGps, defined by a mean and covariance moments, shall be used. These moments depend on a set of hyperparameters which have to be estimated. In the next sections the core formulation that quantifies all these uncertainties using mrGp and Bayes' theorem shall be described.

4.5 Original modular Bayesian approach

4.5.1 General workflow of the modular Bayesian approach

In this section an overview of the multiple response modular Bayesian approach is presented, with a more detailed description given in the following sections. A core formulation was initially developed for a single response case by Kennedy and O'Hagan [134, 135]. Arendt et al. [39] proved that unless under some specific conditions, the single response case will fail to identify the true structural parameters, and proposed instead a multiple response case [146], which allows for a more informative model.

To solve the main equation of model calibration while considering all of the involved uncertainties, the process is broken in four modules, hence the name modular Bayesian approach. For a flowchart of the algorithm see Fig. 4.5. As can be seen, in module 1 the computer model is fitted by a mrGp, based on training simulation data $\boldsymbol{D}^m = \{\boldsymbol{X}^m, \boldsymbol{\theta}^m, \boldsymbol{Y}^m\}$ and the mrGp hyperparameters $\boldsymbol{\phi}^m$ are estimated. Similarly, in module 2 the discrepancy function is fitted by a mrGp, based on training monitored data $\boldsymbol{D}^e = \{\boldsymbol{X}^e, \boldsymbol{Y}^e\}$, the prior of the calibration parameters $p(\boldsymbol{\theta})$ and the model mrGp from module 1, in order to estimate the hyperparameters $\boldsymbol{\phi}^\delta$. Note the superscript m, to differentiate between input simulated data $\boldsymbol{\theta}^m$ and the to-be identified calibration parameters $\boldsymbol{\theta}$. Subsequently, in module 3 the mrGps are used to set up a global data model, the likelihood function $p(\boldsymbol{D}|\boldsymbol{\theta}, \hat{\boldsymbol{\phi}})$, which explains both simulations and observations $\boldsymbol{D} = \{\boldsymbol{D}^m, \boldsymbol{D}^e\}$ for a set of given structural param-

eters $\boldsymbol{\theta}$, and previously determined hyperparameters, now fixed at their estimated values $\hat{\boldsymbol{\phi}} = \{\hat{\boldsymbol{\phi}}^m, \hat{\boldsymbol{\phi}}^\delta\}$. The posterior distribution of the parameters $p(\boldsymbol{\theta}|\boldsymbol{D}, \hat{\boldsymbol{\phi}})$ is estimated by Bayesian inference. The fourth and final module predicts the observed process, by updating the mrGps previously determined, with the posterior information $p(\boldsymbol{\theta}|\boldsymbol{D}, \hat{\boldsymbol{\phi}})$.

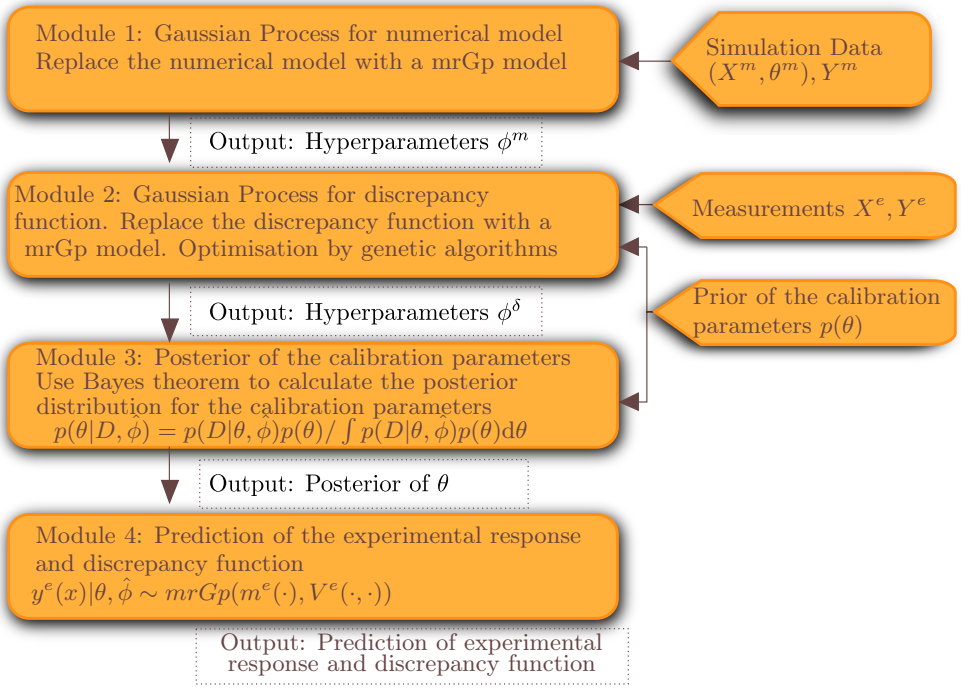


Figure 4.5: Flowchart of the modular Bayesian algorithm, from [39].

One alternative way of determining the hyperparameters is by applying a Bayesian approach which fully accounts for all the considered uncertainties and determines all the hyperparameters at the same time. However this implies a significant computational effort and is not recommended [139]. The MBA fixes the hyperparameters of the mrGp at each of its modules, not fully accounting for existent uncertainties because their 'second order' effect is neglected. This means that preference has been given to recognise all of the present sources of uncertainty, without their interactions and at a lower computational cost, rather than fully accounting for the uncertainties at the cost of an increased computational effort. See Section 4.5 of Kennedy and O'Hagan [134] for additional details.

4.5.2 Module 1 - mrGp of physics-based model

The objective of the MBA first module is to train a mrGP that is fitted to the model response surface, by estimating its hyperparameters $\boldsymbol{\phi}_m$, based on an available simulation dataset $\boldsymbol{D}^m = \{\boldsymbol{X}^m, \boldsymbol{\theta}^m, \boldsymbol{Y}^m\}$. For each row of the input $\{\boldsymbol{X}^m, \boldsymbol{\theta}^m\} \in \mathbb{R}^{N\times d+r}$ there is a corresponding simulated response $\boldsymbol{Y}^m \in \mathbb{R}^{N\times q}$. Note that although the parameters of the model $\boldsymbol{\theta}^m$ are conceptually different from the design variables $\boldsymbol{X}^m$, their treatment when fitting the mrGp is identical.

The procedure described in module 1 is analogous to what has been presented in Section 3.4, where a mrGp is fitted against available training data. The symbols are also analogous to what has been presented before, the only difference is the superscript m, which stands for model.

Summarily, the module can be broken into five distinct steps (cf Fig. 4.6). In Step 1 an assumed mean and correlation function (Sec. 3.2 and 3.3) plus training data D^m are supplied as input. In Step 2 a likelihood function Eq. (4.8) is computed, and then is maximised by genetic algorithms (Step 3) in order to estimate hyperparameters of the mean/correlation functions (Step 4). Finally, Step 5 produces predictions at new unsampled regions using Eq. (4.10) and Eq. (4.11). As shown in Fig. 4.5), the final mrGp maps the model response surface, between the design variables $\boldsymbol{X}$, the calibration parameters $\boldsymbol{\theta}$ and the output responses $\boldsymbol{Y}$, based on the supplied training data.

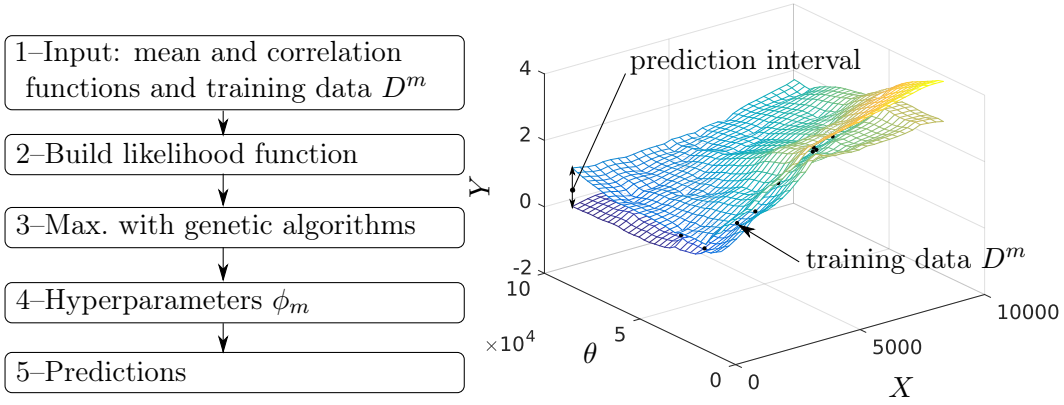


Figure 4.6: Diagram of module 1 – Modular Bayesian approach. Flowchart of multiple steps (left) and illustration of posterior prediction of the model, based on training data (right).

Since the model is deterministic, i.e. the same set of inputs will result in the same output, the mrGp should interpolate perfectly the dataset $\boldsymbol{D}^m$. To enforce

this condition, it suffices to apply the prior structures that have been presented in Section 3.2 and 3.3, with associated hyperparameters $\boldsymbol{\phi}_m = \{\boldsymbol{\beta}_m, \boldsymbol{\Sigma}^2_m, \boldsymbol{\omega}_m\}$ and $\boldsymbol{\omega}_m = \{\boldsymbol{\omega}_x, \boldsymbol{\omega}_\theta\}$, and estimate them by maximising a likelihood function, as described in Section 3.4, with the form

$$p(y|\boldsymbol{\omega}_m, \boldsymbol{D}^m) \propto |\boldsymbol{V}_m|^{-1/2} \tag{4.5}$$

with covariance matrix and process variance estimate

$$\begin{array}{rcl} \boldsymbol{V}_m & = & \hat{\boldsymbol{\Sigma}}^2_m \otimes \boldsymbol{R}_m \\ \hat{\boldsymbol{\Sigma}}^2_m & = & \frac{1}{N}(\boldsymbol{Y}^m - \boldsymbol{H}_m\hat{\boldsymbol{\beta}}_m)^T \boldsymbol{R}_m^{-1}(\boldsymbol{Y}^m - \boldsymbol{H}_m\hat{\boldsymbol{\beta}}_m) \end{array} \tag{4.6}$$

and regression coefficients estimate

$$\begin{array}{rcl} \hat{\boldsymbol{\beta}}_m & = & \boldsymbol{W}_m \boldsymbol{H}_m^T \boldsymbol{R}_m^{-1} \boldsymbol{Y}^m \\ \boldsymbol{W}_m & = & (\boldsymbol{H}_m^T \boldsymbol{R}_m^{-1} \boldsymbol{H}_m)^{-1}. \end{array} \tag{4.7}$$

where $\boldsymbol{W}_m = (\boldsymbol{H}_m^T \boldsymbol{R}_m^{-1} \boldsymbol{H}_m)^{-1} \in \mathbb{R}^{p\times p}$ is the Gram matrix inverse.

An analogous and more tractable expression is the log-likelihood function, denoted as ℓ and given by

$$\ell(y|\boldsymbol{\omega}_m, \boldsymbol{D}^m) \propto -\frac{1}{2}\log|\boldsymbol{V}_m| \tag{4.8}$$

Since the estimates of the process variance matrix $\hat{\boldsymbol{\Sigma}}^2_m$ in Eq. (4.6) and the regression coefficients $\hat{\boldsymbol{\beta}}_m$ in Eq. (4.7) have analytical expressions, the only hyperparameters that have to be determined numerically are the roughness parameters $\boldsymbol{\omega}_m = \{\boldsymbol{\omega}_x, \boldsymbol{\omega}_\theta\}$, included in the correlation matrix $\boldsymbol{R}_m$.

Prediction equations of the Physics-based model mrGp

In this section, the mrGp determined in module 1 will be used to perform predictions at a new set of input points. Also the current section builds directly from what has been presented in Section 3.5, where a full description of the mrGp prediction model has been detailed.

After having observed $\boldsymbol{D}^m$, the posterior distribution of the computer response $\boldsymbol{y}^m(\boldsymbol{x}_*, \boldsymbol{\theta}_*) \in \mathbb{R}^{s\times q}$ given $\boldsymbol{D}^m$, $\boldsymbol{\omega}_m$, $\boldsymbol{\Sigma}^2_m$ and $\boldsymbol{\beta}_m$, is Gaussian with mean and covariance prior functions

$$p(\boldsymbol{y}^m(\boldsymbol{x}_*, \boldsymbol{\theta}_*)) = \mathcal{N}(\boldsymbol{h}_m(\boldsymbol{x}_*, \boldsymbol{\theta}_*)\boldsymbol{\beta}_m, \boldsymbol{\Sigma}^2_m \otimes R_m((\boldsymbol{x}_*, \boldsymbol{\theta}_*), (\boldsymbol{x}'_*, \boldsymbol{\theta_*}'))) \tag{4.9}$$

It is recalled that the matrices R_m and $\boldsymbol{h}_m$ are denoted differently from matrices $\boldsymbol{R}_m$ and $\boldsymbol{H}_m$ just so that it is clear to which dataset each matrix belongs to. Additionally, the arguments $(\boldsymbol{x}_*, \boldsymbol{\theta}_*)$ of these matrices will be omitted for the sake of clarity. The posterior distribution of the model mrGp depends of the hyperparameters, but these can be integrated out and their estimates can be used instead (note the circumflex above $\boldsymbol{\beta}_m$ and $\boldsymbol{\Sigma}^2$). Finally, the expectation of the posterior is

$$\mathrm{E}[\boldsymbol{y}^m|\boldsymbol{D}^m, \hat{\boldsymbol{\phi}}^m] = \boldsymbol{h}_m\hat{\boldsymbol{\beta}}_m + \mathcal{R}_m^T\boldsymbol{R}_m^{-1}(\boldsymbol{Y}^m - \boldsymbol{H}_m\hat{\boldsymbol{\beta}}_m) \tag{4.10}$$

and covariance

$$V[\boldsymbol{y}^m, \boldsymbol{y}^m|\boldsymbol{D}^m, \hat{\boldsymbol{\phi}}^m] = \hat{\boldsymbol{\Sigma}}_m^2 \otimes \{R_m - \mathcal{R}_m^T\boldsymbol{R}_m^{-1}\mathcal{R}_m + \boldsymbol{u}^T[\boldsymbol{H}_m^T\boldsymbol{R}_m^{-1}\boldsymbol{H}_m^{-1}]\boldsymbol{u}\} \tag{4.11}$$

where

$$\boldsymbol{u}(\boldsymbol{x}_*, \boldsymbol{\theta}_*) = \boldsymbol{h}_m(\boldsymbol{x}_*, \boldsymbol{\theta}_*)^T - \boldsymbol{H}_m^T\boldsymbol{R}_m^{-1}\mathcal{R}_m(\boldsymbol{x}_*, \boldsymbol{\theta}_*). \tag{4.12}$$

and $\mathcal{R}_m(\boldsymbol{x}_*, \boldsymbol{\theta}_*) \in \mathbb{R}^{N\times s}$ is a relational correlation matrix, that translates correlations between the new input sites $(\boldsymbol{x}_*, \boldsymbol{\theta}_*)$ and the simulated dataset $(\boldsymbol{X}^m, \boldsymbol{\theta}^m)$, and whose ith row is

$$[R_m((\boldsymbol{x}_{1*}, \boldsymbol{\theta}_{1*}), (\boldsymbol{x}_i^m, \boldsymbol{\theta}_i^m)), R_m((\boldsymbol{x}_{2*}, \boldsymbol{\theta}_{2*}), (\boldsymbol{x}_i^m, \boldsymbol{\theta}_i^m)), \ldots, R_m((\boldsymbol{x}_{s*}, \boldsymbol{\theta}_{s*}), (\boldsymbol{x}_i^m, \boldsymbol{\theta}_i^m))].$$

This predictive model will be used in module 2 of the approach, described in the following section.

4.5.3 Module 2 - mrGp of discrepancy function

In module 2 the objective is, similarly to module 1, to train a mrGp in order to approximate the discrepancy function. However, the current task is considerably more complex than what has been done in module 1. Some remarks are required in order to clarify this approximation.

1. Firstly, the application of the mrGp with observation error as described in Section 3.6 is not direct, since there is an additional term related with the model mrGp.

2. Secondly, it must be noted that the covariance functions for the model, discrepancy function and observation error are assumed *a priori* as statistically independent.

3. Finally, there is the problematic issue that at this stage the structural param-

eters $\boldsymbol{\theta}$ are still unknown.

In order to put the current module into perspective, Fig. 4.7 displays its flowchart, and its five main steps. In Step 1, the calibration parameter prior belief, the training experimental data $\boldsymbol{D}^e = \{\boldsymbol{X}^e, \boldsymbol{Y}^e\}$ and the assumed mean and correlation functions of the discrepancy function mrGp are given as an input.

In Step 2, the prior is numerically integrated into the posterior distribution (in order to remove the dependency of the calibration parameters) of the model mrGp. Note that the posterior is set at input conditions $\boldsymbol{X}^e$. Moreover the model mrGp has fixed hyperparameters $\boldsymbol{\phi}^m$, and its final mean vector is shown in Eq. (4.14), and its covariance matrix in Eq. (4.15).

The remainder steps 3-5 are similar to what has been shown in module 1. A likelihood function is computed, Eq. (4.19), and is maximised with genetic algorithms. The final result of the maximisation are the hyperparameters estimates of the discrepancy function mrGp $\boldsymbol{\phi}_\delta$.

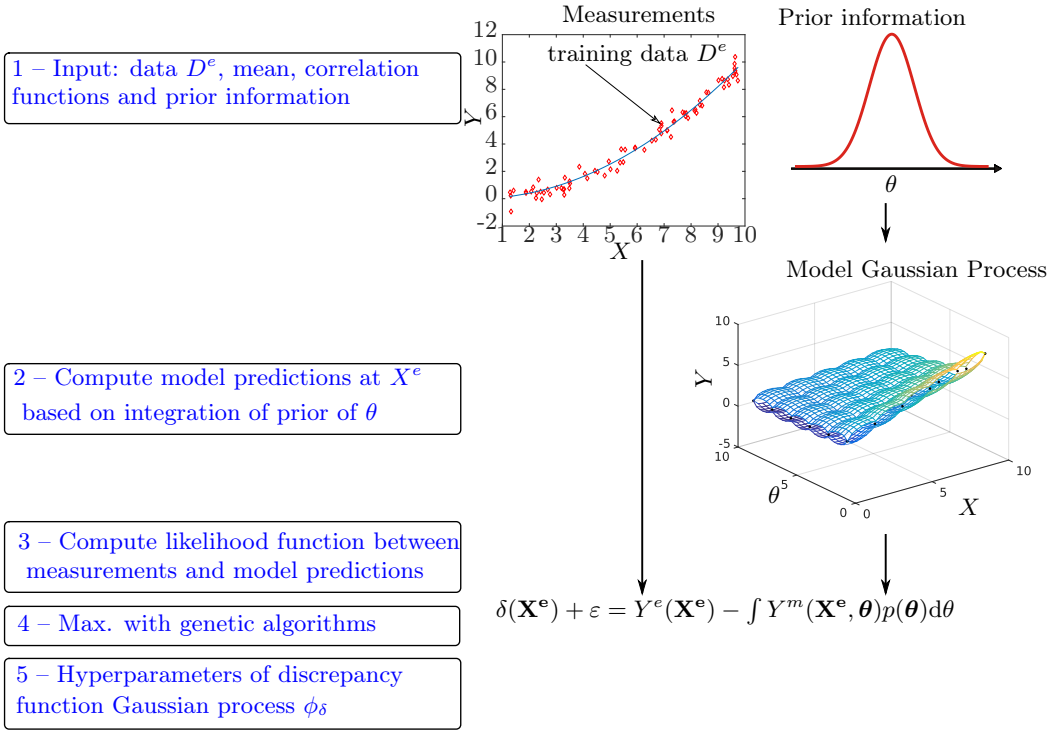


Figure 4.7: Diagram of module 2 workflow – Modular Bayesian approach. Flowchart (left) and correspondent illustration (right). Step 2 integrates model predictions with prior information $p(\theta)$ at X^e. Steps 3-5 are similar to module 1.

Due to the complexity of Step 2, the next section will be devoted exclusively to detail the integration of the model mrGp with the prior PDF of the parameters.

Integration of the model mrGp and prediction at observation sites

As described above, an integration with respect to the prior distribution of the parameters $p(\boldsymbol{\theta})$ is performed over the model mrGp, i.e.

$$\boldsymbol{Y}^e(\boldsymbol{X}^e) \sim \int \mathrm{mrGp}^m(\boldsymbol{X}^e, \boldsymbol{\theta}) p(\boldsymbol{\theta}) \mathrm{d}\boldsymbol{\theta} + \mathrm{mrGp}^\delta(\boldsymbol{X}^e). \tag{4.13}$$

which probability density is multivariate normal

$$p(\boldsymbol{Y}^e | \boldsymbol{D}^m, \hat{\boldsymbol{\phi}}_m, \boldsymbol{\phi}_\delta) \sim \mathcal{N}(\boldsymbol{m}_e, \boldsymbol{V}_e)$$

with mean vector $\boldsymbol{m}_e$ and covariance matrix $\boldsymbol{V}_e$. The expression for the ith row of $\boldsymbol{m}_e$ is derived directly from the sum of Eq. (4.10) (the $\boldsymbol{\theta}$ variable terms are integrated in the $p(\boldsymbol{\theta})$ domain) with the prior mean function of the discrepancy function $\boldsymbol{h}_\delta(x_i^e)\boldsymbol{\beta}_\delta$

$$\begin{aligned} \boldsymbol{m}_{e,i} &= \int \mathrm{E}[\boldsymbol{Y}_i^e | \boldsymbol{D}^m, \hat{\boldsymbol{\phi}}_m, \boldsymbol{\beta}_\delta, \boldsymbol{\theta}] p(\boldsymbol{\theta}) \mathrm{d}\boldsymbol{\theta} \\ &= \int \boldsymbol{h}_m(x_i^e, \boldsymbol{\theta}) p(\boldsymbol{\theta}) \mathrm{d}\boldsymbol{\theta} \hat{\boldsymbol{\beta}}_m + \int \mathcal{R}_m(x_i^e, \boldsymbol{\theta})^T p(\boldsymbol{\theta}) \mathrm{d}\boldsymbol{\theta} \boldsymbol{R}_m^{-1} (\boldsymbol{Y}^m - \boldsymbol{H}_m \hat{\boldsymbol{\beta}}_m) + \boldsymbol{h}_\delta(x_i^e) \boldsymbol{\beta}_\delta. \end{aligned} \tag{4.14}$$

Similarly, the covariance matrix is given by

$$\boldsymbol{V}_e = \int V[\boldsymbol{y}^m(x_i^e, \boldsymbol{\theta}), \boldsymbol{y}^m(x_j^e, \boldsymbol{\theta}) | \boldsymbol{D}^m, \hat{\boldsymbol{\phi}}^m] p(\boldsymbol{\theta}) \mathrm{d}\boldsymbol{\theta} + \boldsymbol{\Sigma}_\delta^2 \otimes \boldsymbol{R}_\delta + \boldsymbol{\Lambda} \otimes \boldsymbol{I}_n \tag{4.15}$$

where $\boldsymbol{I}_n$, $\boldsymbol{\Sigma}_\delta^2$, $\boldsymbol{R}_\delta$, and $\boldsymbol{\Lambda}$, are the identity matrix of dimension n, a discrepancy function process variance matrix, a discrepancy function correlation matrix and an observation error variance matrix, respectively.

The most challenging term of Eq. (4.15) to compute is the integral of the covariance matrix which comes from Eq. (4.11). Fortunately, several of its terms are quadratic, e.g. the $\mathcal{R}_m^T \boldsymbol{R}_m^{-1} \mathcal{R}_m$ term. Because the quadratic form itself is a scalar, and because of a property of invariance of the trace to permutations of the factors in a product, there is a very useful fact which states that

$$x^T A x = tr(x^T A x) = tr(A x x^T) \tag{4.16}$$

for a symmetric matrix A and vector x. It is possible to apply this property to obtain the (i, j) element of the integral in Eq. (4.15), which are developed from the quadratic terms of the covariance posterior distribution of the mrGp model response

in Eq. (4.11), as

$$\int \boldsymbol{R}_m((x_i^e,\boldsymbol{\theta}),(x_j^e,\boldsymbol{\theta}))p(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta} = \int R_m((x_i^e,\boldsymbol{\theta}),(x_j^e,\boldsymbol{\theta}))p(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta} \tag{4.17}$$
$$- tr\left\{\boldsymbol{R}_m^{-1}\int \mathcal{R}_m(x_j^e,\boldsymbol{\theta})\mathcal{R}_m(x_i^e,\boldsymbol{\theta})^T p(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta}\right\}$$
$$+ tr\left\{\boldsymbol{W}_m\int h_m(x_j^e,\boldsymbol{\theta})^T h_m(x_i^e,\boldsymbol{\theta})p(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta}\right\}$$
$$- tr\left\{\boldsymbol{W}_m\boldsymbol{H}_m^T\boldsymbol{R}_m^{-1}\int \mathcal{R}_m(x_j^e,\boldsymbol{\theta})h_m(x_i^e,\boldsymbol{\theta})p(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta}\right\}$$
$$- tr\left\{\boldsymbol{R}_m^{-1}\boldsymbol{H}_m\boldsymbol{W}_m\int h_m(x_j^e,\boldsymbol{\theta})^T\mathcal{R}_m(x_i^e,\boldsymbol{\theta})^T p(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta}\right\}$$
$$+ tr\left\{\boldsymbol{R}_m^{-1}\boldsymbol{H}_m\boldsymbol{W}_m\boldsymbol{H}_m^T\boldsymbol{R}_m^{-1}\int \mathcal{R}_m(x_j^e,\boldsymbol{\theta})\mathcal{R}_m(x_i^e,\boldsymbol{\theta})^T p(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta}\right\},$$

where recalling, the model Gram matrix is $\boldsymbol{W}_m = (\boldsymbol{H}_m^T\boldsymbol{R}_m^{-1}\boldsymbol{H}_m)^{-1}$. Note that Eq. (4.17) depends mostly of the matrices R_m, $\mathcal{R}_m$, $\boldsymbol{R}_m$, $\boldsymbol{H}_m$ which have been presented in module 1. The exception are their arguments, which correspond to the observation sites x_i^e and x_j^e.

Differently from previous implementations of the MBA, in the current thesis the integrals in Eqs. (4.14) and (4.17) were computed with numerical methods. Numerically speaking, this integration is one of the most costly operations of the MBA, particularly for multiple structural parameters $\boldsymbol{\theta}$. Further computational details will be highlighted in Section 4.6.1. Lastly, Section 4.7.3 presents a numerical example to assess the performance of the numerical integration operations under parallel processing.

Finally, after marginalising the model mrGp in order to the prior distribution of the calibration parameters, it is necessary to maximise the likelihood function of the metamodel, that determines the hyperparameters $\boldsymbol{\phi}_\delta = \{\boldsymbol{\Lambda}, \boldsymbol{\beta}_\delta, \boldsymbol{\Sigma}_\delta^2, \boldsymbol{\omega}_\delta\}$.

Approximation of a discrepancy mrGp to experimental data

After computing the integrals of the mean and covariance presented in the previous section, it is possible to determine the discrepancy function hyperparameters by maximisation of the log-likelihood function between observations and the established meta-model. For simplicity, the mean vector will be written as $\boldsymbol{m}_e = \hat{\boldsymbol{Y}}^m + \boldsymbol{H}_\delta\boldsymbol{\beta}_\delta$. Writing now, similarly to Eq. (3.10), the multivariate normal likelihood distribution

with the above two moments results in

$$p(\boldsymbol{Y}^e|\boldsymbol{D}^m,\hat{\boldsymbol{\phi}}_m,\boldsymbol{\phi}_\delta) \propto |\boldsymbol{V}_e|^{-1/2} \times \exp\left\{-\frac{1}{2}\text{vec}(\boldsymbol{Y}^e-\hat{\boldsymbol{Y}}^m-\boldsymbol{H}_\delta\hat{\boldsymbol{\beta}}_\delta)^T\boldsymbol{V}_e^{-1}\text{vec}(\boldsymbol{Y}^e-\hat{\boldsymbol{Y}}^m-\boldsymbol{H}_\delta\hat{\boldsymbol{\beta}}_\delta)\right\}, \tag{4.18}$$

where $\text{vec}(\cdot)$ is the vectorising operation. Alternatively a more tractable and equivalent expression of the log-likelihood is given by

$$\ell(\boldsymbol{Y}^e|\boldsymbol{D}^m,\hat{\boldsymbol{\phi}}_m,\boldsymbol{\phi}_\delta) \propto -\frac{1}{2}\{\log|\boldsymbol{V}_e| + \text{vec}(\boldsymbol{Y}^e-\hat{\boldsymbol{Y}}^m-\boldsymbol{H}_\delta\hat{\boldsymbol{\beta}}_\delta)^T \boldsymbol{V}_e^{-1}\text{vec}(\boldsymbol{Y}^e-\hat{\boldsymbol{Y}}^m-\boldsymbol{H}_\delta\hat{\boldsymbol{\beta}}_\delta)\}, \tag{4.19}$$

with MLE of the regression coefficients $\boldsymbol{\beta}_\delta$ as

$$\hat{\boldsymbol{\beta}}_\delta = \boldsymbol{W}_\delta I_q \otimes \boldsymbol{H}_\delta^T\boldsymbol{V}_e^{-1}\text{vec}(\boldsymbol{Y}^e-\hat{\boldsymbol{Y}}^m) \tag{4.20}$$

$$\boldsymbol{W}_\delta = (I_q \otimes \boldsymbol{H}_\delta^T\boldsymbol{V}_e^{-1}I_q \otimes \boldsymbol{H}_\delta)^{-1}. \tag{4.21}$$

Therefore, this objective function has to be maximised in order to estimate the remaining hyperparameters $\boldsymbol{\omega}_\delta$, $\boldsymbol{\Lambda}$ and $\boldsymbol{\Sigma}_\delta^2$. Due to the complexity of this operation, numerical optimisation methods such as genetic algorithms have to be used. Additional details can be found on page 179.

Note that to improve the computational efficiency of the algorithm all of the data has been standardised, and assumed as belonging to a standard multivariate normal distribution. See Appendix C.2 for additional details.

4.5.4 Module 3 - Bayes' theorem

Both of the previous modules where targeted at building a global data model, i.e. the likelihood function, that explains both simulations and observations. These datasets have been denoted as matrices $\boldsymbol{Y}^m \in \mathbb{R}^{N\times q}$ and $\boldsymbol{Y}^e \in \mathbb{R}^{n\times q}$ for simulations and observations, respectively. A combined data vector and dataset shall now be written as

$$\boldsymbol{Y} = \begin{bmatrix}\text{vec}(\boldsymbol{Y}^m)\\ \text{vec}(\boldsymbol{Y}^e)\end{bmatrix} \qquad \boldsymbol{D} = \{\boldsymbol{D}^m, \boldsymbol{D}^e\}, \tag{4.22}$$

respectively. On the other hand, the hyperparameters estimated previously are denoted as $\boldsymbol{\phi} = \{\boldsymbol{\phi}_m, \boldsymbol{\phi}_\delta\}$.

Similarly to the previous two modules, Fig. 4.8 illustrates module 3 workflow into five distinct steps. In Step 1 the hyperparameters obtained in the previous modules $\boldsymbol{\phi}$, plus the whole data set $\boldsymbol{D}$ and the prior information of the parameters $p(\boldsymbol{\theta})$

is supplied. In Step 2, a global mrGp data model which explains both simulations and observations is built. This model is based on the hyperparameters determined in the previous modules which are now fixed, and its mean and covariance are given by Eqs. (4.27) and (4.28), respectively. Step 3 computes the likelihood between the model of Step 2 and the global dataset for different $\boldsymbol{\theta}$ values. Finally, in Step 4 and 5 the likelihood function and the prior information are combined using Bayes' Theorem, Eq. (4.23), in order to obtain the posterior distribution.

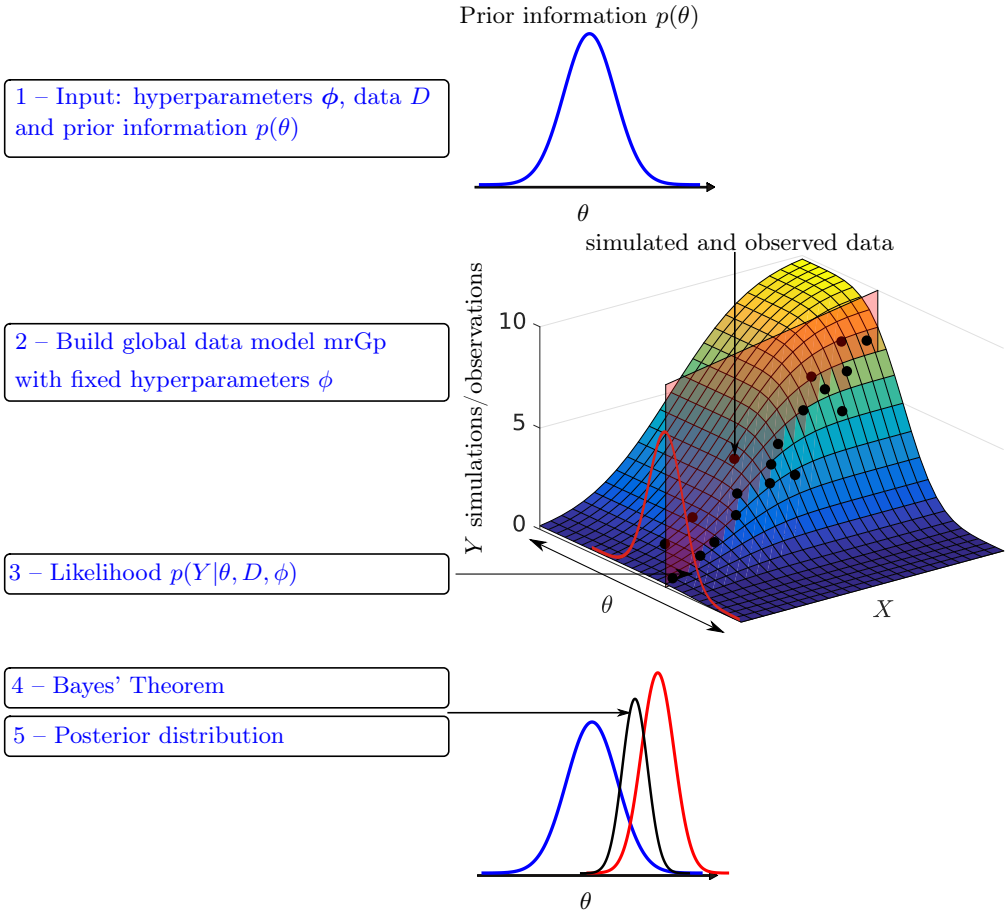


Figure 4.8: Diagram of module 3 workflow – Modular Bayesian approach. Its flowchart (left) and its illustration (right) highlight how the prior information and the data/hyperparameters used in previous modules are used to make inference of the true parameters θ^* using Bayes' theorem.

On the basis of the procedure described above, Bayes' theorem can be re-stated as follows

$$p(\boldsymbol{\theta}|\boldsymbol{Y}, \boldsymbol{D}, \hat{\boldsymbol{\phi}}) = \frac{p(\boldsymbol{Y}|\boldsymbol{\theta}, \boldsymbol{D}, \hat{\boldsymbol{\phi}})p(\boldsymbol{\theta})}{p(\boldsymbol{Y}|\boldsymbol{D}, \hat{\boldsymbol{\phi}})}, \tag{4.23}$$

with marginal likelihood which can be calculated by the following integral

$$p(\boldsymbol{Y}|\boldsymbol{D},\hat{\phi}) = \int p(\boldsymbol{Y}|\boldsymbol{\theta},\boldsymbol{D},\hat{\phi})p(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta}. \tag{4.24}$$

Next, the likelihood function in Eq. (4.23) shall be detailed. Due to the assumptions of the metamodel developed in the previous sections, the likelihood function follows a multivariate normal distribution

$$p(\boldsymbol{Y}|\boldsymbol{\theta},\boldsymbol{D},\hat{\phi}) \sim \mathcal{N}(\boldsymbol{m}_Y(\boldsymbol{\theta}),\boldsymbol{V}_Y(\boldsymbol{\theta})). \tag{4.25}$$

whose mean vector is

$$\boldsymbol{m}_Y(\boldsymbol{\theta}) = \mathrm{E}[\boldsymbol{Y}|\boldsymbol{\theta},\boldsymbol{D},\hat{\phi}] = \boldsymbol{H}_Y(\boldsymbol{\theta})\hat{\boldsymbol{\beta}}, \tag{4.26}$$

where, $\boldsymbol{H}_Y(\boldsymbol{\theta})\hat{\boldsymbol{\beta}}$ are a combination of the linear regression functions of the model and discrepancy function, defined as

$$\boldsymbol{H}_Y(\boldsymbol{\theta})\hat{\boldsymbol{\beta}} = \begin{bmatrix} \boldsymbol{I}_q \otimes \boldsymbol{H}_m(\boldsymbol{X}^m,\boldsymbol{\theta}^m) & \boldsymbol{O} \\ \boldsymbol{I}_q \otimes \boldsymbol{H}_m(\boldsymbol{X}^e,\boldsymbol{\theta}) & \boldsymbol{I}_q \otimes \boldsymbol{H}_\delta(\boldsymbol{X}^e) \end{bmatrix} \begin{bmatrix} \mathrm{vec}(\hat{\boldsymbol{\beta}}_m) \\ \mathrm{vec}(\hat{\boldsymbol{\beta}}_\delta) \end{bmatrix}, \tag{4.27}$$

where $\boldsymbol{O}$ is a matrix of all zeros. Similarly the covariance matrix is defined as

$$\boldsymbol{V}_Y(\boldsymbol{\theta}) = V[\boldsymbol{Y}|\boldsymbol{\theta},\boldsymbol{D},\hat{\phi}] \tag{4.28}$$

and follows the form presented in Eq. (3.7) to represent the covariance of the simulations and observations (in compact form) as

$$\begin{bmatrix} \hat{\boldsymbol{\Sigma}}_m^2 \otimes \boldsymbol{R}_m(D_m) & \hat{\boldsymbol{\Sigma}}_m^2 \otimes \boldsymbol{R}_m(D_m, D_e(\boldsymbol{\theta})) \\ \hat{\boldsymbol{\Sigma}}_m^2 \otimes \boldsymbol{R}_m(D_m, D_e(\boldsymbol{\theta}))^T & \hat{\boldsymbol{\Lambda}} \otimes \boldsymbol{I}_n + \hat{\boldsymbol{\Sigma}}_m^2 \otimes \boldsymbol{R}_m(D_e(\boldsymbol{\theta})) + \hat{\boldsymbol{\Sigma}}_\delta^2 \otimes \boldsymbol{R}_\delta(D_e) \end{bmatrix}. \tag{4.29}$$

Note however that the estimated regression coefficients should now be determined from this global data model as

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{H}_Y(\boldsymbol{\theta})^T\boldsymbol{V}_Y(\boldsymbol{\theta})^{-1}\boldsymbol{H}_Y(\boldsymbol{\theta}))^{-1}\boldsymbol{H}_Y(\boldsymbol{\theta})^T\boldsymbol{V}_Y(\boldsymbol{\theta})^{-1}\boldsymbol{Y},$$

instead of using the hyperparameters estimated in the previous modules. By plugging these moments into the multivariate normal density function, the posterior

distribution is obtained as

$$p(\boldsymbol{\theta}|\boldsymbol{Y},\boldsymbol{D},\hat{\boldsymbol{\phi}}) \propto p(\boldsymbol{\theta})|\boldsymbol{V}_Y(\boldsymbol{\theta})|^{-1/2}|\boldsymbol{W}_Y(\boldsymbol{\theta})|^{1/2} \times \exp\left\{-\frac{1}{2}(\boldsymbol{Y}-\boldsymbol{H}_Y(\boldsymbol{\theta})\hat{\boldsymbol{\beta}})^T\boldsymbol{V}_Y(\boldsymbol{\theta})^{-1}(\boldsymbol{Y}-\boldsymbol{H}_Y(\boldsymbol{\theta})\hat{\boldsymbol{\beta}})\right\}. \quad (4.30)$$

Note that this PDF also depends of the determinant of a Gramian matrix $\boldsymbol{W}_Y$ because of the integration of the regression coefficients $\boldsymbol{\beta}$ out of the density. The proof is lengthy and therefore it has been moved to Appendix B. The other hyperparameters have been assumed to have a non-informative prior, and therefore do not affect the above equation.

The PDF described above is used to make inference regarding the true value of the parameters $\boldsymbol{\theta}^*$. Assuming that the data under consideration is sufficiently informative in respect to the whereabouts of $\boldsymbol{\theta}^*$, the discrepancy function will also effectively highlight the inadequacy of the model. The posterior PDF of $\boldsymbol{\theta}$ has to be used in the fourth and last module of the methodology to compute the unconditional prediction of the true process.

Due to its complexity, the integration of the marginal likelihood in Eq. (4.24) has been performed numerically. For lower order examples (one structural parameter) the Gauss-Legendre algorithm can be used, otherwise it is necessary to resort to MCMC methods. The current implementation supports both of these methods. This enhancement shall be more thoroughly detailed in Section 4.6.2.

4.5.5 Module 4 - model prediction

Finally, having determined the posterior distribution of the structural parameters $\boldsymbol{\theta}^*$ and hyperparameters $\boldsymbol{\phi}$ on the previous modules, it is now possible to predict the response at unobserved points $\boldsymbol{x}_*$.

It is important to mention that, up to this point the predictions of a mrGp have always been presented as conditional on input data, e.g. as given by Eq. (3.25) where the function $\boldsymbol{y}$ depends of some input sites $\boldsymbol{x}^*$. Although a similar structure is applied in this module it is necessary to marginalise out the posterior distribution determined in module 3. In other words, the response *conditional* on the posterior $p(\boldsymbol{\theta}|\boldsymbol{Y},\boldsymbol{D},\hat{\boldsymbol{\phi}})$ has to be transformed with the law of total expectation and variance, in order to retrieve a response that does not depend on $\boldsymbol{\theta}$ [197], i.e. into an *unconditional* response. The reason why these laws have to be used, is to enclose the estimation uncertainty of the posterior on the predicted responses.

In summary, the law of total expectation states that two random variables coexistent in the same probability space, X and Y, can relate their conditional and

unconditional expected value as follows

$$\mathrm{E}(Y) = \mathrm{E}(\mathrm{E}(Y|X)) \tag{4.31}$$

where $\mathrm{E}(Y)$ is the expected value of Y. Similarly, the law of total variance states that

$$V(Y) = \mathrm{E}[V(Y|X)] + V(\mathrm{E}[Y|X]) \tag{4.32}$$

as long as the variance of Y is finite.

The application of the above laws and formulation of the unconditional predicted model, experiment and discrepancy function responses shall be detailed separately in the following sections.

Physics-based model

A final illustration of the main steps required to compute a model response prediction is shown in Fig. 4.9. Essentially, in Step 1 the posterior PDF of $\boldsymbol{\theta}$, hyperparameters $\boldsymbol{\phi}$ and the points $\boldsymbol{x}_*$ where prediction of the q responses will take place, are supplied as an input. Subsequently, in Step 2 the posterior of the mrGp conditional on the posterior of the parameters $\boldsymbol{\theta}$ is established. Finally, in Step 3 and 4, the unconditional response of the mrGp, which removes the dependency of $\boldsymbol{\theta}$ from the posterior mrGp, is computed through the law of total expectation and variance, resulting in Eqs. (4.33) and (4.34). As it shall be demonstrated in the next sections, the same principles have to be applied to predict the unconditional experimental response and discrepancy function.

It is relatively straightforward to generate unconditional predictions with the physics-based model, given that its conditional response has already been presented in Section 4.5.2.

The two moments of the *conditional* model response at a given point $(\boldsymbol{x}_*, \boldsymbol{\theta})$ are analogous to Eqs. (4.10) and (4.11) and are denoted as $\mathrm{E}[\boldsymbol{y}^m(\boldsymbol{x}_*)|\boldsymbol{\theta}, \boldsymbol{D}^m, \hat{\boldsymbol{\phi}}^m]$ and $\boldsymbol{V}[\boldsymbol{y}^m(\boldsymbol{x}_*), \boldsymbol{y}^m(\boldsymbol{x}'_*)|\boldsymbol{\theta}, \boldsymbol{D}^m, \hat{\boldsymbol{\phi}}^m]$ for the mean vector and covariance matrix, respectively. Note that the only difference is that the calibration parameter $\boldsymbol{\theta}$ was removed from the function argument and was placed after the vertical bar. In other words $\boldsymbol{\theta}$ is now considered as a random variable upon which the response is conditioned, rather than as an input data point. As mentioned before, calculating the unconditional posterior distribution requires applying the law of total expectation and variance as follows

$$\mathrm{E}[\boldsymbol{y}^m(\boldsymbol{x}_*)|\boldsymbol{D}^m, \hat{\boldsymbol{\phi}}^m] = \mathrm{E}[\mathrm{E}[\boldsymbol{y}^m(\boldsymbol{x}_*)|\boldsymbol{\theta}, \boldsymbol{D}^m, \hat{\boldsymbol{\phi}}^m]] \tag{4.33}$$

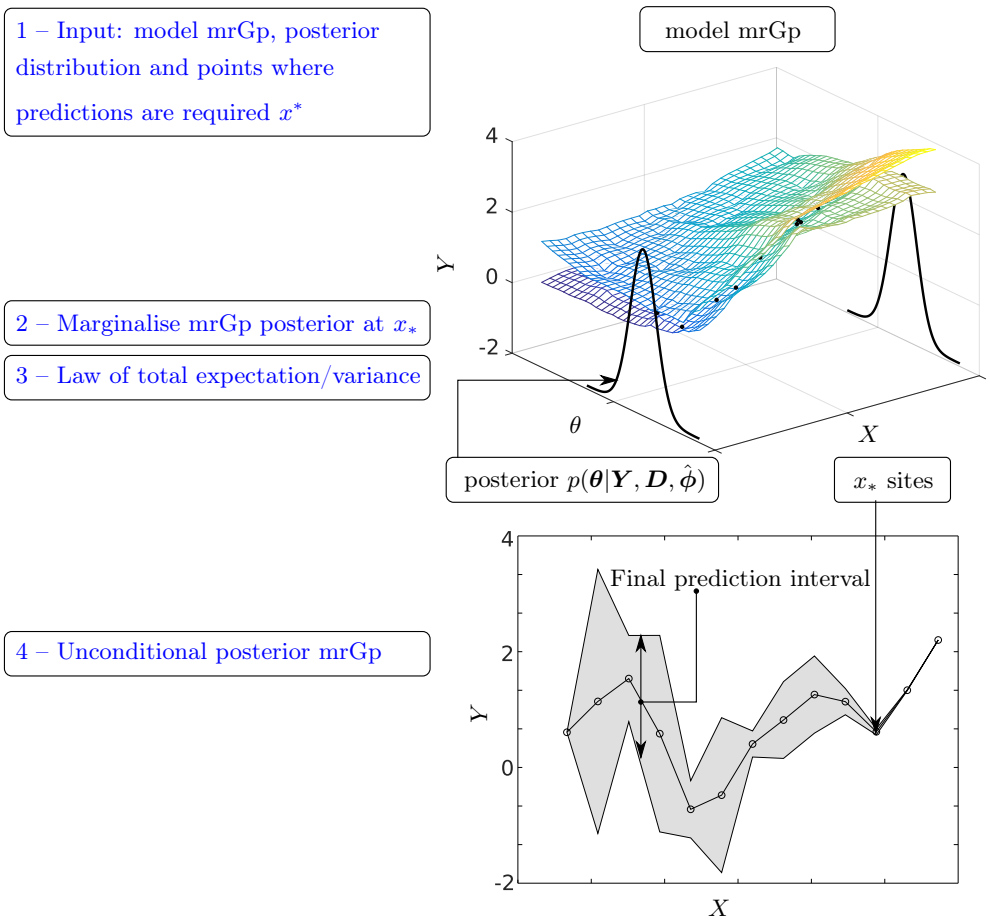


Figure 4.9: Diagram of module 4 for prediction of the unconditional physics-based model – Modular Bayesian approach. The flowchart (left) and the illustration (right) display how the posterior of θ and the model mrGp can be used to compute unconditional predictions at x^*.

for the mean vector, and

$$\begin{aligned} V[\boldsymbol{y}^m(\boldsymbol{x}_*)\boldsymbol{y}^m(\boldsymbol{x}'_*)|\boldsymbol{D}^m, \hat{\boldsymbol{\phi}}^m] &= \mathrm{E}[V[\boldsymbol{y}^m(\boldsymbol{x}_*), \boldsymbol{y}^m(\boldsymbol{x}'_*)|\boldsymbol{\theta}, \boldsymbol{D}^m, \hat{\boldsymbol{\phi}}^m]] \\ &+ V[\mathrm{E}[\boldsymbol{y}^m(\boldsymbol{x}_*)|\boldsymbol{\theta}, \boldsymbol{D}^m, \hat{\boldsymbol{\phi}}^m], \mathrm{E}[\boldsymbol{y}^m(\boldsymbol{x}'_*)|\boldsymbol{\theta}, \boldsymbol{D}^m, \hat{\boldsymbol{\phi}}^m]] \end{aligned} \tag{4.34}$$

for the covariance matrix. In practice, the process requires the Gauss-Legendre probability densities or the Metropolis Hastings samples of the parameters posterior.

Experimental response

On the other hand, the experimental response is a more complex response and requires a more careful detailing than the physics-based model. The two moments of

the *conditional* experimental response are derived from the metamodel which was built in module 1 and module 2, and are given by

$$\mathrm{E}[\boldsymbol{y}^e(\boldsymbol{x}_*)^T|\boldsymbol{\theta},\boldsymbol{D},\hat{\boldsymbol{\phi}}] = \boldsymbol{h}_Y(\boldsymbol{x}_*,\boldsymbol{\theta})\hat{\boldsymbol{\beta}} + \mathcal{V}(\boldsymbol{x}_*,\boldsymbol{\theta})^T\boldsymbol{V}_Y(\boldsymbol{\theta})^{-1}(\boldsymbol{Y} - \boldsymbol{H}_Y(\boldsymbol{\theta})\hat{\boldsymbol{\beta}}) \tag{4.35}$$

for the mean vector, and

$$\begin{aligned}\boldsymbol{V}[\boldsymbol{y}^e(\boldsymbol{x}_*)^T,\boldsymbol{y}^e(\boldsymbol{x}'_*)^T|\boldsymbol{\theta},\boldsymbol{D},\hat{\boldsymbol{\phi}}] &= \hat{\boldsymbol{\Sigma}}^2_m \otimes R_m((\boldsymbol{x}_*,\boldsymbol{\theta}),(\boldsymbol{x}'_*,\boldsymbol{\theta})) + \hat{\boldsymbol{\Sigma}}^2_\delta \otimes R_\delta(\boldsymbol{x}_*,\boldsymbol{x}'_*) \\ &+ \hat{\boldsymbol{\Lambda}} \otimes \boldsymbol{I}_s - \mathcal{V}(\boldsymbol{x}_*,\boldsymbol{\theta})^T\boldsymbol{V}_Y(\boldsymbol{\theta})^{-1}\mathcal{V}(\boldsymbol{x}'_*,\boldsymbol{\theta}) \\ &+ \boldsymbol{U}(\boldsymbol{x}_*,\boldsymbol{\theta})^T(\boldsymbol{H}_Y(\boldsymbol{\theta})^T\boldsymbol{V}_Y(\boldsymbol{\theta})^{-1}\boldsymbol{H}_Y(\boldsymbol{\theta}))^{-1}\boldsymbol{U}(\boldsymbol{x}'_*,\boldsymbol{\theta})\end{aligned} \tag{4.36}$$

where the combined relational covariance matrix is defined as

$$\mathcal{V}(\boldsymbol{x}_*,\boldsymbol{\theta}) = \begin{bmatrix} \hat{\boldsymbol{\Sigma}}^2_m \otimes \mathcal{R}_m((\boldsymbol{X}^m,\boldsymbol{\Theta}^m),(\boldsymbol{x}_*,\boldsymbol{\theta})) \\ \hat{\boldsymbol{\Sigma}}^2_m \otimes \mathcal{R}_m((\boldsymbol{X}^e,\boldsymbol{\theta}),(\boldsymbol{x}_*,\boldsymbol{\theta})) + \hat{\boldsymbol{\Sigma}}^2_\delta \otimes \mathcal{R}_\delta(\boldsymbol{X}^e,\boldsymbol{x}_*) \end{bmatrix}, \tag{4.37}$$

and

$$\boldsymbol{h}_Y(\boldsymbol{x}_*,\boldsymbol{\theta}) = [\boldsymbol{I}_q \otimes \boldsymbol{h}_m(\boldsymbol{x}_*,\boldsymbol{\theta}) \quad \boldsymbol{I}_q \otimes \boldsymbol{h}_\delta(\boldsymbol{x}_*)] \tag{4.38}$$

$$\boldsymbol{U}(\boldsymbol{x}_*,\boldsymbol{\theta}) = \boldsymbol{h}_Y(\boldsymbol{x}_*,\boldsymbol{\theta})^T - \boldsymbol{H}_Y(\boldsymbol{\theta})^T\boldsymbol{V}_Y(\boldsymbol{\theta})^{-1}\mathcal{V}(\boldsymbol{x}_*,\boldsymbol{\theta}) \tag{4.39}$$

for the covariance matrix. Section 3.6 highlights the details of the posterior model which has been used herein. Notice also that in these equations the global covariance and regression matrices $\boldsymbol{V}_Y$ and $\boldsymbol{H}_Y$, assembled in module 3 of the formulation (Section 4.5.4), are required.

After assembling these moments, the unconditional response is obtained in a similar manner as to what has been presented in the previous section, applying the law of total expectation and total variance.

$$\mathrm{E}[\boldsymbol{y}^e(\boldsymbol{x}_*)^T|\boldsymbol{D},\hat{\boldsymbol{\phi}}] = \mathrm{E}[\mathrm{E}[\boldsymbol{y}^e(\boldsymbol{x}_*)^T|\boldsymbol{\theta},\boldsymbol{D},\hat{\boldsymbol{\phi}}]] \tag{4.40}$$

$$\begin{aligned}V[\boldsymbol{y}^e(\boldsymbol{x}_*)^T\boldsymbol{y}^e(\boldsymbol{x}'_*)^T|\boldsymbol{D},\hat{\boldsymbol{\phi}}] &= \mathrm{E}[V[\boldsymbol{y}^e(\boldsymbol{x}_*)^T\boldsymbol{y}^e(\boldsymbol{x}'_*)^T|\boldsymbol{\theta},\boldsymbol{D},\hat{\boldsymbol{\phi}}]] \\ &+ V[\mathrm{E}[\boldsymbol{y}^e(\boldsymbol{x}_*)^T|\boldsymbol{\theta},\boldsymbol{D},\hat{\boldsymbol{\phi}}],\mathrm{E}[\boldsymbol{y}^e(\boldsymbol{x}'_*)^T|\boldsymbol{\theta},\boldsymbol{D},\hat{\boldsymbol{\phi}}]]\end{aligned} \tag{4.41}$$

Due to the complexity of the posterior distribution of the structural parameters this marginalisation requires numerical integration methods.

Discrepancy function

The two moments of the conditional discrepancy function are given by

$$\begin{aligned} \mathrm{E}[\boldsymbol{\delta}(\boldsymbol{x}_*)|\boldsymbol{\theta}, \boldsymbol{D}, \hat{\boldsymbol{\phi}}] &= \boldsymbol{h}_\delta(\boldsymbol{x}_*)\mathrm{vec}(\hat{\boldsymbol{\beta}^\delta}) + \mathcal{R}_\delta(\boldsymbol{x}_*)^T \\ &\times \boldsymbol{R}_\delta^{-1}\mathrm{vec}(\boldsymbol{Y}^e - \hat{\boldsymbol{Y}}^m - \boldsymbol{H}_\delta\hat{\boldsymbol{\beta}}^\delta) \end{aligned} \tag{4.42}$$

for its expected value and by

$$\begin{aligned} \boldsymbol{V}[\boldsymbol{\delta}(\boldsymbol{x}_*), \boldsymbol{\delta}(\boldsymbol{x}'_*)|\boldsymbol{\theta}, \boldsymbol{D}, \hat{\boldsymbol{\phi}}] = \quad & \hat{\boldsymbol{\Sigma}}_\delta^2 \otimes \{R_\delta(\boldsymbol{x}_*, \boldsymbol{x}'_*) - \mathcal{R}_\delta(\boldsymbol{x}_*)^T\boldsymbol{R}_\delta^{-1}\mathcal{R}_\delta(\boldsymbol{x}'_*) \\ & +(\boldsymbol{h}_\delta(\boldsymbol{x}_*)^T - \boldsymbol{H}_\delta^T\boldsymbol{R}_\delta^{-1}\mathcal{R}_\delta(\boldsymbol{x}_*))^T(\boldsymbol{H}_\delta^T\boldsymbol{R}_\delta^{-1}\boldsymbol{H}_\delta)^{-1}(\boldsymbol{h}_\delta(\boldsymbol{x}'_*)^T - \boldsymbol{H}_\delta^T\boldsymbol{R}_\delta^{-1}\mathcal{R}_\delta(\boldsymbol{x}'_*))\} \end{aligned} \tag{4.43}$$

for its covariance, where

$$\mathcal{R}_\delta(\boldsymbol{x}_*) = \mathcal{R}_\delta(\boldsymbol{X}^e, \boldsymbol{x}_*) \tag{4.44}$$

However for simplicity and computational reasons the discrepancy function expected value shall be calculated simply as

$$\mathrm{E}[\boldsymbol{\delta}(\boldsymbol{x}_*)|\boldsymbol{D}, \hat{\boldsymbol{\phi}^\delta}] = \mathrm{E}[\boldsymbol{y}^e(\boldsymbol{x}_*)|\boldsymbol{D}, \hat{\boldsymbol{\phi}}] - \mathrm{E}[\boldsymbol{y}^m(\boldsymbol{x}_*)|\boldsymbol{D}^m, \hat{\boldsymbol{\phi}^m}] \tag{4.45}$$

After having determined these moments, it is possible to obtain an associated prediction interval with a specific percentage error and represent it along with the mean function. For the current work an interval of $m \pm 2\sqrt{V}$, i.e. a 95% confidence interval was assumed. These results will be illustrated in the numerical examples in Section 4.7.

This section concludes the presentation of the modular Bayesian approach for multiple responses. In the next sections enhancements over this formulation shall be highlighted and discussed.

4.6 Enhanced modular Bayesian approach

The previous sections have detailed the general formulation of the MBA, inspired by the work of Kennedy and O'Hagan [134] and Arendt et al. [146]. In the current section, a number of enhancements which are novel contributions to the applicability of the MBA to SHM will be presented. Specifically, the MBA has been expanded to consider identification of multiple calibration parameters.

4.6.1 Numerical integration of the observation mrGp - module 2

As detailed in Section 4.5.3, the discrepancy function $\boldsymbol{\delta}$ is approximated by a mrGp in the second module of the MBA. The task involves three steps:

1. integration of the calibration parameters prior information onto the model mrGp determined in module 1;

2. generation of new predictions at observed input sites;

3. and maximisation of a likelihood function given observed data and the built metamodel.

The main focus of the current section is step 1, the integration of the calibration parameters. As shown in previous works, this operation was always computed for a single parameter with closed-form solutions, specifically for Gaussian and uniform priors. Instead, the author implemented the routine numerically, which allows:

- to consider any type of prior distribution;

- to expand the MBA for multiple parameter calibration.

A MATLAB module has been developed jointly with a package for multi-dimensional Markov chain integration [198], which computes the integrals of Eqs. (4.14) and (4.15). The implemented package allows to control the absolute or relative error associated with the integration, which by default has been set at a 1×10^{-2} relative error.

Although in general the results of the implemented routine have been positive, it should also be noted that an inappropriately selected model correlation function (see Table 3.2 and surrounding text) can result in numerical instabilities, causing, e.g., an integrated covariance matrix which is not positive semi-definite.

Finally, it should be mentioned that each entry of the covariance matrix that is being integrated can be computed independently of the others, which makes it particularly suitable for parallel computing. An example which highlights the algorithm performance will be illustrated in Section 4.7.3.

4.6.2 MCMC sampling of posterior distribution - module 3

Along with the module discussed in the previous section, some additional enhancements are required in module 3 of the MBA in order to allow identification of multiple calibration parameters.

Bayesian inference is used to determine the posterior distribution of $\boldsymbol{\theta}$, and the integral in Eq. (4.24) is the marginalised posterior distribution, which previous authors have determined with Gauss-Legendre quadrature.

Now, to expand the MBA for multiple structural parameters, it is necessary to perform the computation using MCMC methods. Specifically, the Metropolis–Hastings (MH) algorithm [142, 143] has been chosen. Hence, the aforementioned integral can be approximated from a large number of samples S, which are drawn from a designated target distribution $q(\boldsymbol{\theta})$

$$\int f(\boldsymbol{\theta}) q(\boldsymbol{\theta}) \mathrm{d}\boldsymbol{\theta} \approx \frac{1}{S} \sum_{t=1}^{S} f(\boldsymbol{\theta}^{(t)}). \tag{4.46}$$

where $f(\boldsymbol{\theta})$ is a given function. In the context of this work $f(\boldsymbol{\theta})$ and $q(\boldsymbol{\theta})$ correspond to the prior and likelihood PDF, respectively. One inherent aspect to the MH algorithm is that a proposal distribution, as similar to the target distribution as possible, has to be specified. However, since it is known that the likelihood is multivariate normal and the data has been standardised, it is relatively safe to assume the proposal distribution has a standard multivariate normal distribution, i.e. $\mathcal{N}(\boldsymbol{O}, \boldsymbol{I})$.

The algorithm generates a Markov chain of samples, which are at each generation step accepted or rejected according to the ratio between the previous and current probability density of a sample, symmetry of the target distribution, and other considerations. The computational effort of the MH can also be parallelised by considering multiple Markov chains.

The MH is known to converge to the target distribution for an increasing number of samples. However it is very important to ensure that there is no correlation between accepted samples, factor which can be assessed by examining a trace plot of the samples.

Although it is acknowledged that the MH algorithm has its limitations, and an alternative such as the adaptive Metropolis algorithm would be more suitable [199], the aim is to showcase the potential of the MBA for multiple parameter inference and motivate further developments. In Chapter 5 the algorithm performance shall be illustrated with an application to the Tamar long suspension bridge.

With the above described enhancements, the MBA is designed to be computationally efficient, even when considering several calibration parameters and multiple responses. To the author's best knowledge this is the first implementation of the MBA formulation which allows multiple responses and multiple calibration parameters simultaneously. In the next sections numerical examples that illustrate its

performance shall be highlighted.

4.7 Numerical examples

To illustrate the algorithm performance, a set of numerical examples have been developed and tested. These test the capability of the algorithm to handle single or multiple design variables and structural parameters, outliers, different types of prior information, noise and residual variations. The problem is based on a cantilever beam subjected to a point end load and a point bending moment. Observations/ simulations which will be used for uncertainty quantification include the beam tip displacement, rotation angle, and deformation energy.

4.7.1 Single design variable and structural parameter

In this example a cantilever beam subjected to a point end load suffers a deflection, which is being observed/simulated at the end of the beam. Additionally the deformation energy of the beam is also considered as an output response. The discrepancy between the model and the actual cantilever beam occurs because of a rotational spring (of stiffness K^*) located at the beginning of the beam, and the parameter that has to be identified based on the observations is the Young's modulus of the beam E^*, as shown in Fig. 4.10.

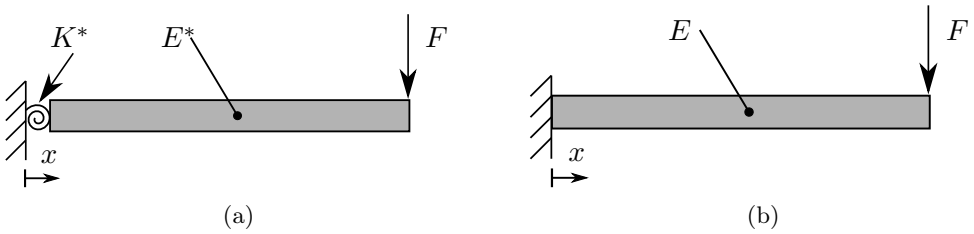


Figure 4.10: Cantilever beam example. Actual (a) and idealised cantilever beam (b).

Numerical values of the beam properties are shown in Table 4.1, including sampling intervals for all of the involved parameters.

Parameter	Numerical value	Parameter	Numerical value
K^*	15×10^{11} Nmm/rad	L	3000 mm
E^*	70×10^3 MPa	E	$[20, 100] \times 10^3$ MPa
F	$[1, 10] \times 10^3$ N	I	6.75×10^8 mm^4
A	300×300 mm^2		

Table 4.1: Parameters of the cantilever beam.

To discriminate the equations of the model and experiment under consideration, it suffices to apply normal beam theory to determine the equations of deflection and deformation energy. For the physics-based model this will be

$$y^m = \frac{FL^3}{3EI} \tag{4.47}$$

for the beam end displacement, and

$$U^m = \frac{F^2L^3}{6EI} \tag{4.48}$$

for the deformation energy. For the actual beam (experimental, hence the e superscript) we need to consider the effect of the rotational spring in addition to the previous equations, i.e.

$$y^e = \frac{FL^3}{3E^*I} + \frac{FL^2}{K^*}, \tag{4.49}$$

for the beam end displacement, and

$$U^e = \frac{F^2L^3}{6E^*I} + \frac{F^2L^2}{K^*} \tag{4.50}$$

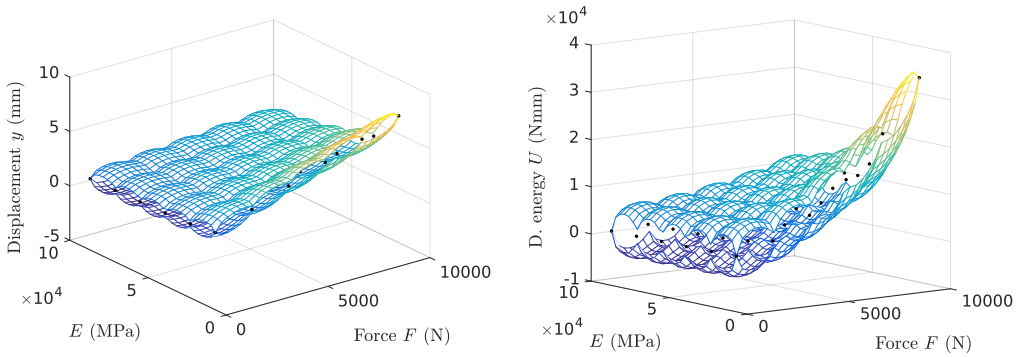
for the deformation energy. By comparison of these equations, it is evident that the second terms of the actual beam equations are the discrepancy function (dependency of K^*).

Firstly, a simulation dataset dependent on force F and Young's modulus E is run by the physics-based model on a $[F, E]$ 6×6 grid, totalling $N = 36$ points. An mrGp has been fitted to the training dataset and is shown in Fig. 4.11. Similarly to previously shown plots of mrGp, the 3D probabilistic model is represented by a cloud of uncertainty, which shrinks near training data points and expands when distant. If there were more design variables or calibration parameters the same type of interpolation of uncertainty would apply, although its visualisation would be less straightforward. For the mrGp shown in Fig. 4.11 its hyperparameters are

$$\hat{\boldsymbol{\omega}}_{F,M,E} = \begin{bmatrix} 0.128 & 0.119 \end{bmatrix} \hat{\boldsymbol{\beta}}_{y,U} = \begin{bmatrix} 0.381 & 0.588 \end{bmatrix} \quad \hat{\boldsymbol{\Sigma}}^2 = \begin{bmatrix} 3.738 & 0 \\ 0 & 124841038 \end{bmatrix}$$

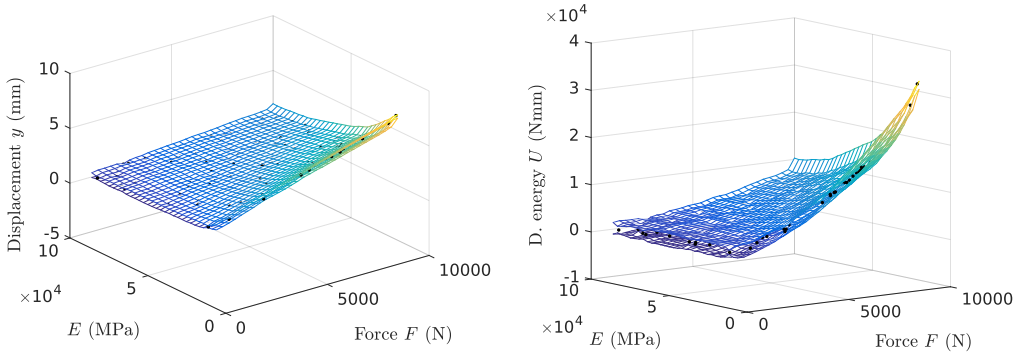
Alternatively, the mrGp can be trained with data which has been generated in a Latin Hypercube sample configuration [200, 201] as shown in Fig. 4.12. In general the latter provides better results when compared with the mesh grid, and has been used in the following examples by default.

Secondly, the discrepancy function has also been fitted by a mrGp, which



(a) (b)

Figure 4.11: Probabilistic predictions of the cantilever beam model based on a mrGp trained with a simulated dataset of 36 samples (black points) in a 6x6 grid configuration. MrGp prediction of displacement at beam tip (a) and of deformation energy (b).



(a) (b)

Figure 4.12: Probabilistic predictions of the cantilever beam model based on a mrGp trained with a simulated dataset of 70 samples (black dots) in a Latin-hypercube configuration. MrGp prediction of displacement at beam tip (a) and of deformation energy (b).

similarly to the model mrGp, also has been assumed to have a linear correlation function and a constant regression polynomial function $\boldsymbol{H}(\cdot) = 1$. See Tables 3.1 and 3.2 for clarification. To fully describe the used mrGps it is necessary to state their hyperparameters. For the mrGp fitted to the model in the Latin hypercube configuration these are

$$\hat{\boldsymbol{\omega}}_{F,E} = [0.179 \quad 0.145] \qquad \hat{\boldsymbol{\Sigma}} = \begin{bmatrix} 0.427 & 0.0 \\ 0.0 & 1.5 \times 10^7 \end{bmatrix}$$
$$\hat{\boldsymbol{\beta}}_{y,U} = [0.709 \quad 0.968]$$

and for the discrepancy function mrGp

$$\hat{\boldsymbol{\omega}}_F = 0.024 \qquad \hat{\boldsymbol{\Sigma}} = \begin{bmatrix} 0.012 & 0.0 \\ 0.0 & 8.5 \times 10^6 \end{bmatrix}.$$
$$\hat{\boldsymbol{\beta}}_{y,U} = [1.5284 \quad 0.8106]$$

Note that the observation error variance matrix $\boldsymbol{\Lambda}$ is not presented, since for this example observation error has not been considered. Moreover, the process variance $\hat{\boldsymbol{\Sigma}}$ is much smaller than for the mesh grid example shown above.

Thirdly, inference on E^* takes place in the third module of the MBA, based on the mrGps approximated in the previous modules. Prior information of the Young's modulus was set as a uniform distribution in the $[20, 100] \times 10^3$ MPa range, i.e. as an uninformative prior. The estimated likelihood and posterior PDF are shown in Fig. 4.13. The moments of the posterior distribution are $\mathrm{E}[\theta] = 70302$ MPa and $V[\theta] = 6915^2$ MPa^2, which result in a relative error of $\varepsilon_r = 0.43\%$, against the reference true value. It should be noted that in this figure the randomness around the mean represents an uncertainty associated with the estimation process, which in a simple example as currently shown would decrease with an increase of input data. At the limit the PDF would converge to a Dirac delta function.

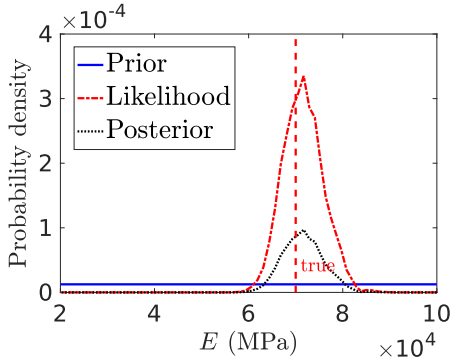


Figure 4.13: PDF representing inference of Young's modulus parameter and its true discrete value (vertical dashed line) for 25 observation points.

Fourthly, the simulated response of the improved metamodel is shown in Fig. 4.14. These plots were computed by marginalising the mrGps of the model and discrepancy function with the posterior information of the calibration parameters, as described in Section 4.5.5. The true model discrepancy/experimental response in diamond markers were computed from the last terms of Eqs. (4.49) and (4.50). It is visible that the prediction interval (PI) of the model discrepancy tends to be larger in the region where there is a larger force applied, which is reasonable due to the fact that the variability of the output is also larger in this region. As can

be seen, the last point exhibits an offset, which occurs due to the lack of simulated data in this region, visible also in Figs. 4.12(a) and 4.12(b). However, the prediction interval at 95 % was able to capture this uncertainty effectively. Note also that the PI's near the regions where training data is missing (cf. Figs. 4.14(e) and 4.14(f) at the 3–5 kN interval) become larger and follow the same trend as the process, i.e. a linear adjustment for the prediction of the displacement and a more rounded shape for the deformation energy.

Finally, the influence of the number of simulated and observed points on the expected value and standard deviation of the identified Young's modulus E was tested. An increasing number of data points was supplied as input, ranging from a minimum of 5 to 300 points. Results can be seen in Fig. 4.15 and Fig. 4.16. From these plots it can be seen that an increase in the number of simulated points does not negatively affect the performance of the approach. However for a relatively large number of data points, 265 points and upwards, the computational effort was considerable, the algorithm started developing numerical instability issues, and was unable to recover a solution consistently. This limitation is partly related to the chosen kernel and the size and condition number of the mrGp covariance matrices.

Relatively to the number of observations, Fig. 4.16(b) indicates that after a certain amount of data points (140) the standard deviation increases rapidly. From 215 points onwards the algorithm is unable to consistently recover a solution, mainly because of the model discrepancy induced by the rotational spring. Therefore, it is important to note that an excessive amount of observed data might effectively decrease the precision of the identification process. Obviously as the complexity of the problem increases, e.g. with additional structural parameters, observation error, observed responses and design variables (as shown in the following sections) the required amount of data will also increase significantly. It is also important to ensure that the input data contains as much information relative to the operational and environmental conditions that the structural system is subject to. This will ensure that the emulated model is capable of generalising predictions beyond the observed data and does not remain over-fitted in a specific region.

4.7.2 Multiple design variables and single structural parameter

Similarly to the previous section, a cantilever beam is now subjected to a point end load and an additional end moment. Its deflection is observed at the end of the beam and the deformation energy is also considered. The current example illustrates the ability of the algorithm to approximate multi-dimensional processes with several input features. The difference between the model and the actual cantilever beam is

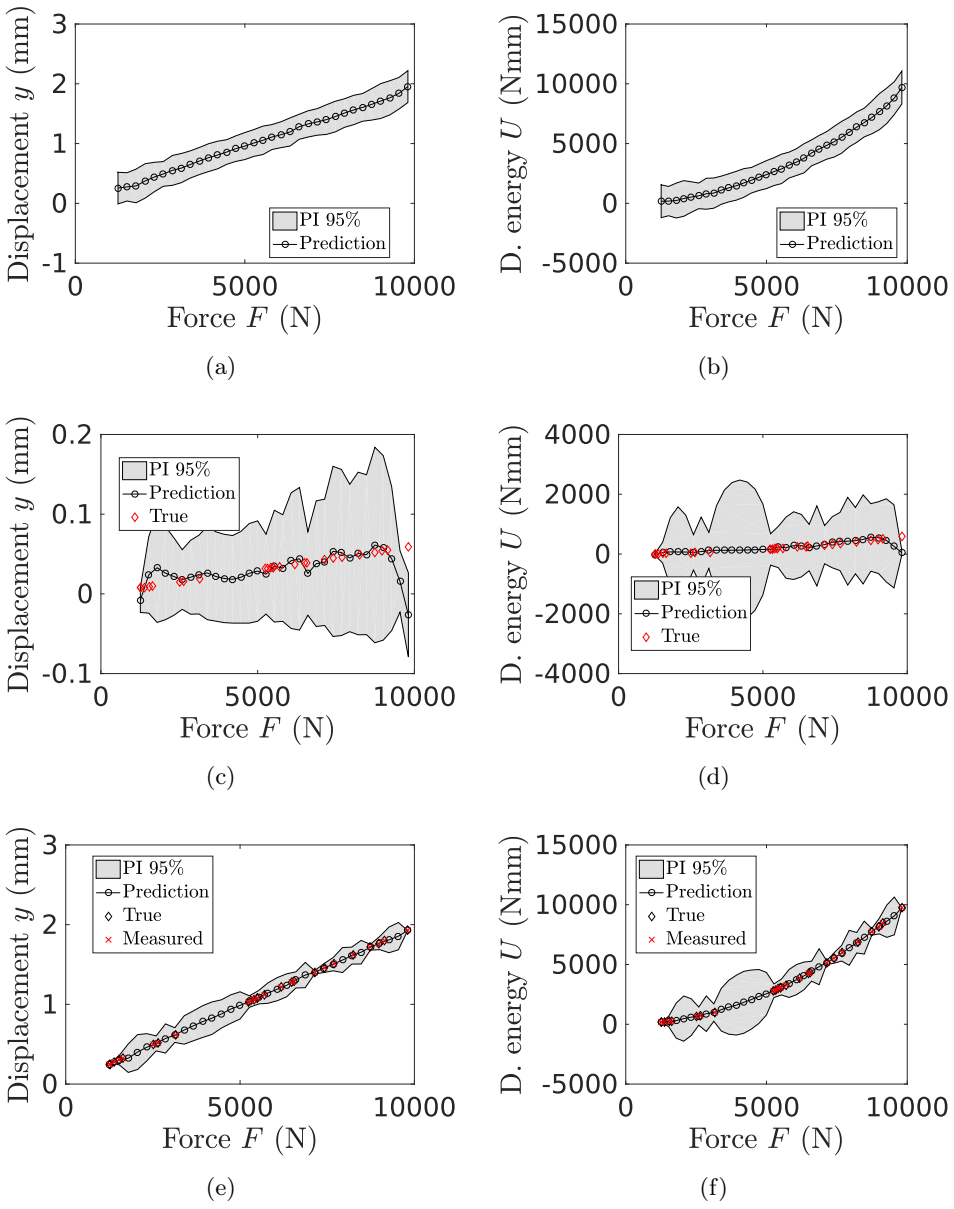
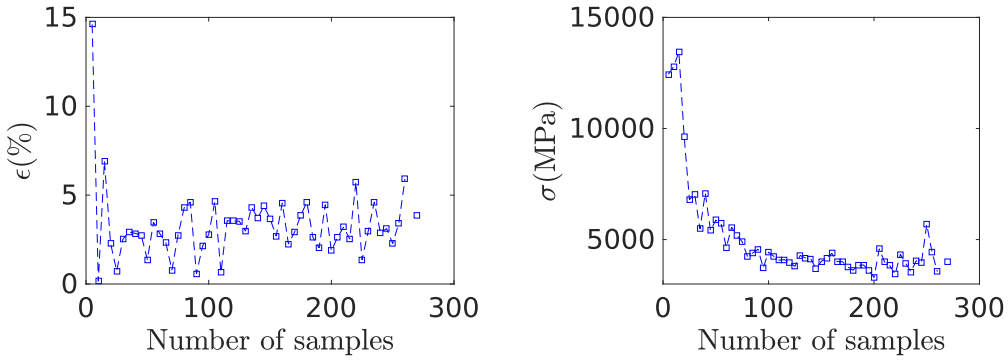


Figure 4.14: Predictions of the cantilever beam model (a), (b), discrepancy function (c) (d) and final predictions (e) (f). True discrepancy function and experimental response represented as diamond markers.

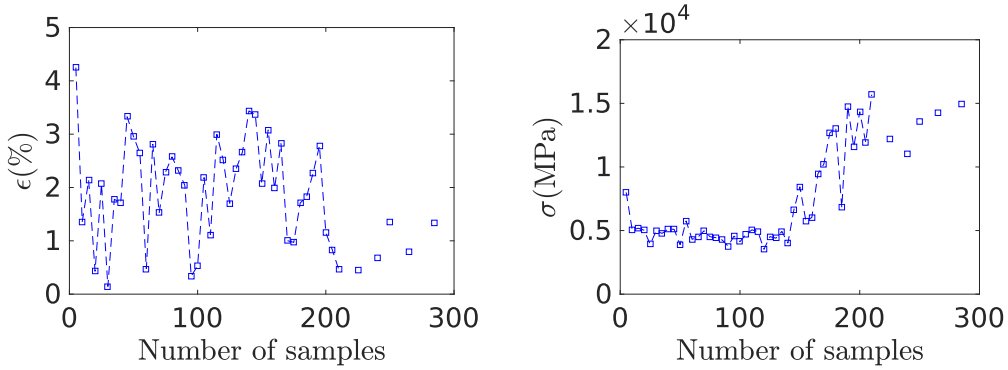
a rotational spring K^* located at the beginning of the beam; and the parameter that has to be identified based on the observations is the Young's modulus of the beam E^*, as shown in Fig. 4.17.

Numerical values of the beam properties are shown in Tables 4.1 and 4.2.



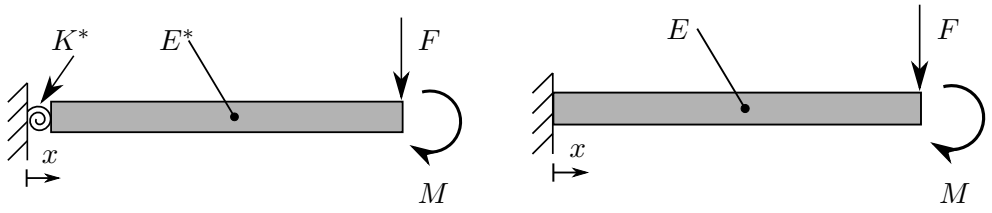
(a) (b)

Figure 4.15: Effect of number of simulated data points for a fixed value of 25 observed data points. Relative error (a) and standard deviation (b) of the cantilever beam Young's modulus.



(a) (b)

Figure 4.16: Effect of number of observed data points for a fixed value of 80 simulated data points. Relative error (a) and standard deviation (b) of the cantilever beam Young's modulus.



(a) (b)

Figure 4.17: Cantilever beam example with multiple design variables. Actual (a) and idealised cantilever beam (b).

Parameter	Numerical value
M	$[1, 30] \times 10^6$ Nmm

Table 4.2: Properties of the cantilever beam.

From normal beam theory the deflection equation for the model is

$$y^m = \frac{FL^3}{3EI} + \frac{ML^2}{2EI},$$

and the deformation energy equation is given by

$$U^m = \frac{F^2L^3}{6EI} + \frac{M^2L}{2EI} \tag{4.51}$$

On the other hand the actual beam displacement is given by

$$y^e = \frac{FL^3}{3E^*I} + \frac{ML^2}{2E^*I} + \frac{FL^2}{K^*} + \frac{ML}{K^*},$$

and for the deformation energy it is given by

$$U^e = \frac{F^2L^3}{6E^*I} + \frac{M^2L}{2E^*I} + \frac{F^2L^2}{K^*} + \frac{M^2}{K^*}. \tag{4.52}$$

If the model and actual beam equations are compared, it can be seen that the terms dependent on K^* represent the discrepancy between the simulations and the experimental data. Although this is a very simple example, it allows to test the MBA under several challenging scenarios. More situations will be highlighted in the next sections.

Firstly, the mrGp which fits the model response surface cannot be plotted, since it is a four $y^m(F, M, E)$ dimensional function. However its hyperparameters are

$$\hat{\boldsymbol{\omega}}_{F,M,E} = \begin{bmatrix} 0.021 & 0.078 & 0.166 \end{bmatrix} \; \hat{\boldsymbol{\beta}}_{y,U} = \begin{bmatrix} 0.752 & 0.934 \end{bmatrix}$$

$$\hat{\boldsymbol{\Sigma}}^2 = \begin{bmatrix} 3.303 & 0 \\ 0 & 234382899.698 \end{bmatrix},$$

and for the discrepancy function mrGp the hyperparameters are

$$\hat{\boldsymbol{\omega}}_{F,M} = \begin{bmatrix} 0.002 & 0.006 \end{bmatrix} \quad \hat{\boldsymbol{\beta}}_{y,U} = \begin{bmatrix} -0.378 & -0.329 \end{bmatrix}$$

$$\hat{\boldsymbol{\Sigma}}^2 = \begin{bmatrix} 0.037 & 0 \\ 0 & 719245138.812 \end{bmatrix}.$$

Secondly, the inference of the structural parameter (module 3 of the MBA) for 150 simulated data points and 66 observations is shown in Fig. 4.18. It is possible to perceive how much data is required for each additional design variable or structural parameter, by contrasting these numbers with the 70 and 25 simulated/observed data points of the previous section. Such numbers correspond to doubling the amount of input data for one additional design variable. Lastly, the moments of the posterior distribution are $\mathrm{E}[\theta] = 69862$ MPa and $V[\theta] = 6127^2$ MPa^2, which results in a relative error of $\varepsilon_r = 0.20\%$, against the reference true value.

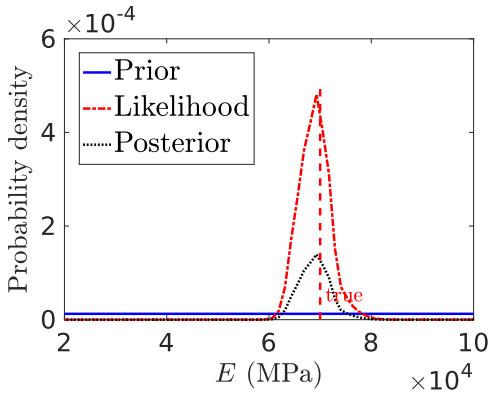


Figure 4.18: Inference of Young's modulus stiffness parameter for a dataset of 150 simulation and 66 observation points.

Thirdly, the predictions from module 4 of the MBA of the discrepancy function and the experimental response are shown in Fig. 4.19. Oppositely to the previous section, the probabilistic models are now 3D surfaces, with the force and bending moment in the abscissas and the deformation energy and displacement in the ordinates. As a consequence of the accurate identification of the Young's modulus it can be seen that the predicted discrepancy function was able to capture the true values of model discrepancy, as observed in plots in Fig. 4.19(a) and 4.19(b). However the number of observations should be carefully considered, ideally by examining the evolution of the standard deviation, as shown in the previous section in Fig. 4.15 and Fig. 4.16. Secondly, due to the reduced amount of data that usually populate the corners of the sampled space there is more uncertainty in these regions, in a very similar manner as what has already been shown in Figs. 4.14(f) and 4.14(d) of the previous section. Lastly, the final predictions of the experimental response, in Figs. 4.19(c) and 4.19(d), also interpolate perfectly the experimental training data, since no observation error has been considered.

To conclude the current example, it is important to mention that in contrast

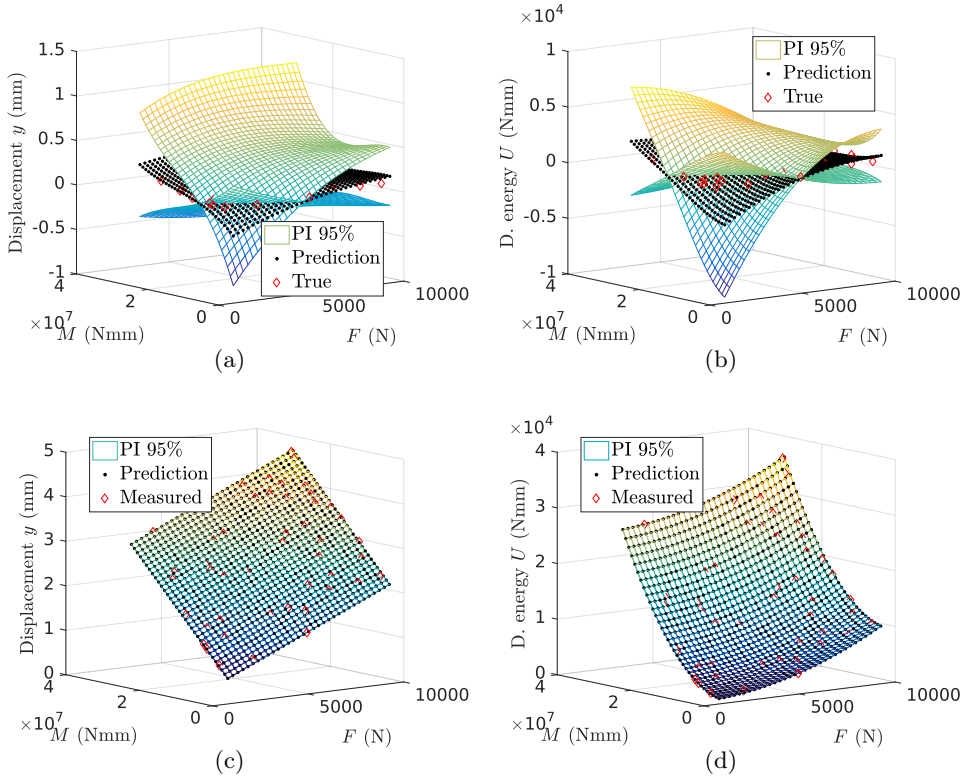


Figure 4.19: Predictions of the cantilever beam discrepancy function (a) (b), and of experiment by metamodel (c) (d). True discrepancy function and experimental response represented as diamond markers.

with the hierarchical Bayes framework, the MBA allows to consider environmental/ operational effects through the design variables $\boldsymbol{X}$ (the force and moment applied on beam's tip for this example). Although it is preferable for the design variables to be controllable, e.g. a in a forced vibration modal identification test, environmental factors such as temperature, humidity, wind can also be considered. The only main requirement is that the variable is directly measurable in the monitored structure. Finally, it would be interesting to compare the ability of the hierarchical framework and the MBA to represent these external effects.

4.7.3 Single design variable and multiple structural parameters

In this numerical example the efficiency of the algorithm to estimate multiple structural parameters is tested. This feature is an enhancement over the original formulation, which depends upon the Metropolis–Hastings algorithm, which samples the likelihood distribution as described in Section 4.6.2. The current implementation

of the MBA is the first one which is capable of considering multiple responses and identify multiple parameters simultaneously.

The same simulated example of a cantilever beam is used. Fig. 4.20(a) represents the actual structure, with a flexible support and true values of Young's modulus E^* and cross sectional inertia I^*. On the other hand Fig. 4.20(b) is the model version of the beam, with a fixed support and two unknown structural parameters. Numer-

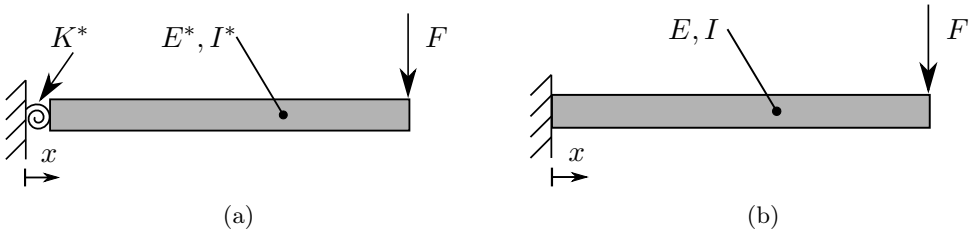


Figure 4.20: Cantilever beam example with multiple parameters inference: Young's modulus and cross sectional inertia. Actual (a) and idealised cantilever beam (b).

ical values of the beam properties are shown in Table 4.3. To make inference of the beam Young's modulus and the cross sectional inertia, these quantities are assumed to have uniform priors with intervals in the specified region.

Parameter	Numerical value	Parameter	Numerical value
E^*	70×10^3 MPa	E	$[20, 100] \times 10^3$ MPa
I^*	6.75×10^8 mm^4	I	$[2.25, 11.25] \times 10^8$ mm^4

Table 4.3: Parameters of the cantilever beam.

Two responses will be considered to aid the inference process: the end displacement of the beam and its deformation energy. The formulas for these quantities are available from normal beam theory as

$$y^m = \frac{FL^3}{3EI} \qquad y^e = \frac{FL^3}{3E^*I^*} + \frac{FL^2}{K^*}$$

for the displacement and

$$U^m = \frac{F^2L^3}{6EI} \qquad U^e = \frac{F^2L^3}{6E^*I^*} + \frac{1}{2}\frac{F^2L^2}{K^*}$$

for the deformation energy. The m and e superscript stand for model and experiment, respectively. By comparison of these equations it is evident that the second term in the experimental quantities is the discrepancy function (dependency of K^*). Formally, the input given to the modular Bayesian approach is a matrix

$[\boldsymbol{X}^m, \Theta^m] = [F, E, I]$ combined with $\boldsymbol{Y}^m = [y^m, U^m]$ for the model and $\boldsymbol{X}^e = [F]$ and $\boldsymbol{Y}^e = [y^e, U^e]$ for the experiment. The intervals that these variables take are displayed in Table 4.1 and 4.3.

At first, modules 1 and 2 of the MBA will estimate the hyperparameters of mrGp which are fitted to the displacement and deformation energy responses. Similarly to what has been shown in the previous sections, a linear correlation function and a zero-order polynomial regression function have been set in the mean and covariance functions. The estimated hyperparameters are

$$\hat{\boldsymbol{\omega}}_{F,E,I} = \begin{bmatrix} 0.184 & 0.156 & 0.157 \end{bmatrix} \; \hat{\boldsymbol{\beta}}_{y,U} = \begin{bmatrix} 0.826 & 0.97 \end{bmatrix}$$

$$\hat{\boldsymbol{\Sigma}}^2 = \begin{bmatrix} 1.116 & 0 \\ 0 & 23657034.464 \end{bmatrix}$$

for the model mrGp, and

$$\hat{\boldsymbol{\omega}}_F = 0.184 \quad \hat{\boldsymbol{\beta}}_{y,U} = \begin{bmatrix} -0.413 & -0.428 \end{bmatrix}$$

$$\hat{\boldsymbol{\Sigma}}^2 = \begin{bmatrix} 0.012 & 0 \\ 0 & 20775434.858 \end{bmatrix}$$

for the discrepancy function mrGp, and were obtained with 170 simulated points and 70 observations, which is a considerably larger amount of data than what has been used in the previous examples.

In contrast, module 3 of the MBA requires additional considerations in order to sample the posterior distribution of the two parameters using the Metropolis Hastings algorithm. For this example 40 000 samples were computed, with 1000 samples selected as an initial burn-in period. This burn-in period aims at reducing the effect of low density regions, that might occur in the initial phase of the convergence process towards the target distribution. Four markov chains were used to obtain the current results, with their acceptance rates [36.65 36.34 36.57 34.31]%. The trace plot of the four chains is shown in Fig. 4.21, and no correlation between the sampled points is visible. It is important to observe the plot, in order to ensure that the Markov chains have indeed memorylessness [202]. Fig. 4.22 and Table 4.4 display the main results of the inference of the structural parameters., and as can be seen, the algorithm is able to infer both parameters with an error within $\epsilon \leq 4.88\%$ range.

The large amount of data which was required for this inference is justified by the fact that a uniform prior was assumed. In Bayesian statistics, such assumption often requires a large amount of data, and the solution is not regularised, as it would be

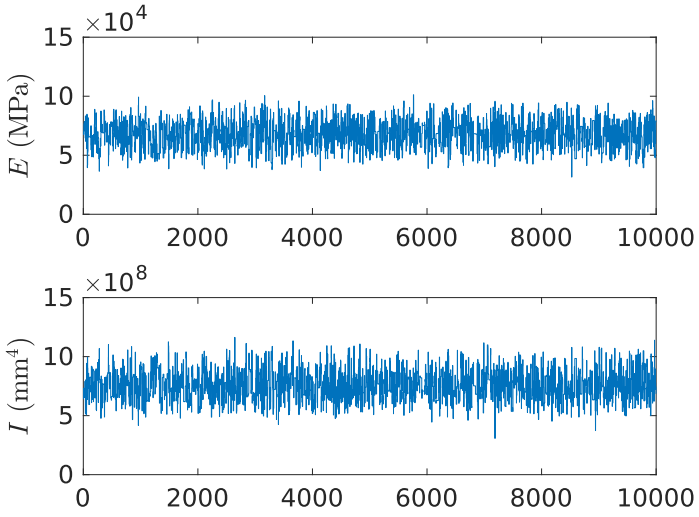


Figure 4.21: Trace plot of samples in the Markov chain.

Parameter θ	E$[\theta]$	$\sigma[\theta]$	ϵ (%)
$\hat{E}$	67.906×10^3 MPa	11.067×10^3 MPa	2.99
$\hat{I}$	7.08×10^8 mm^4	1.25×10^8 mm^4	4.88

Table 4.4: Inference of Young's modulus and inertia.

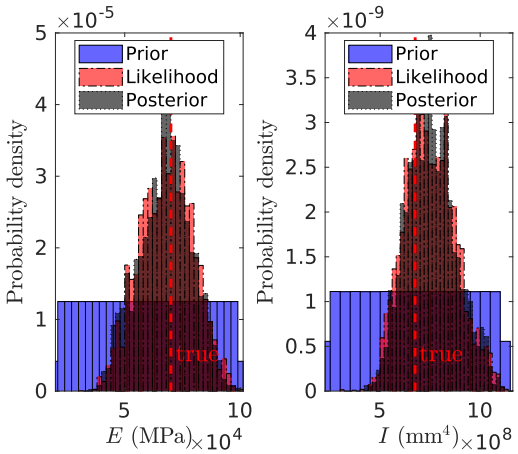


Figure 4.22: Projection of bivariate posterior distribution and reference true values (dashed line).

for a more informative prior. Moreover, the implemented sampling routine is prone to develop correlated Markov chains if the likelihood is not well defined, i.e. was built with sufficient data and/or enough responses.

Similarly to the previous sections, the predictions of the Gaussian process

metamodels are shown in Fig. 4.23. Although not obvious, the operations required to obtain these predictions are different from the ones previously presented. Specifically, the marginalisations required by the law of total expectation and total variance in module 4 had previously been computed with the Gauss-Legendre quadrature, while in this example these have been obtained with Monte Carlo integration, using the samples from the posterior PDF obtained in module 3.

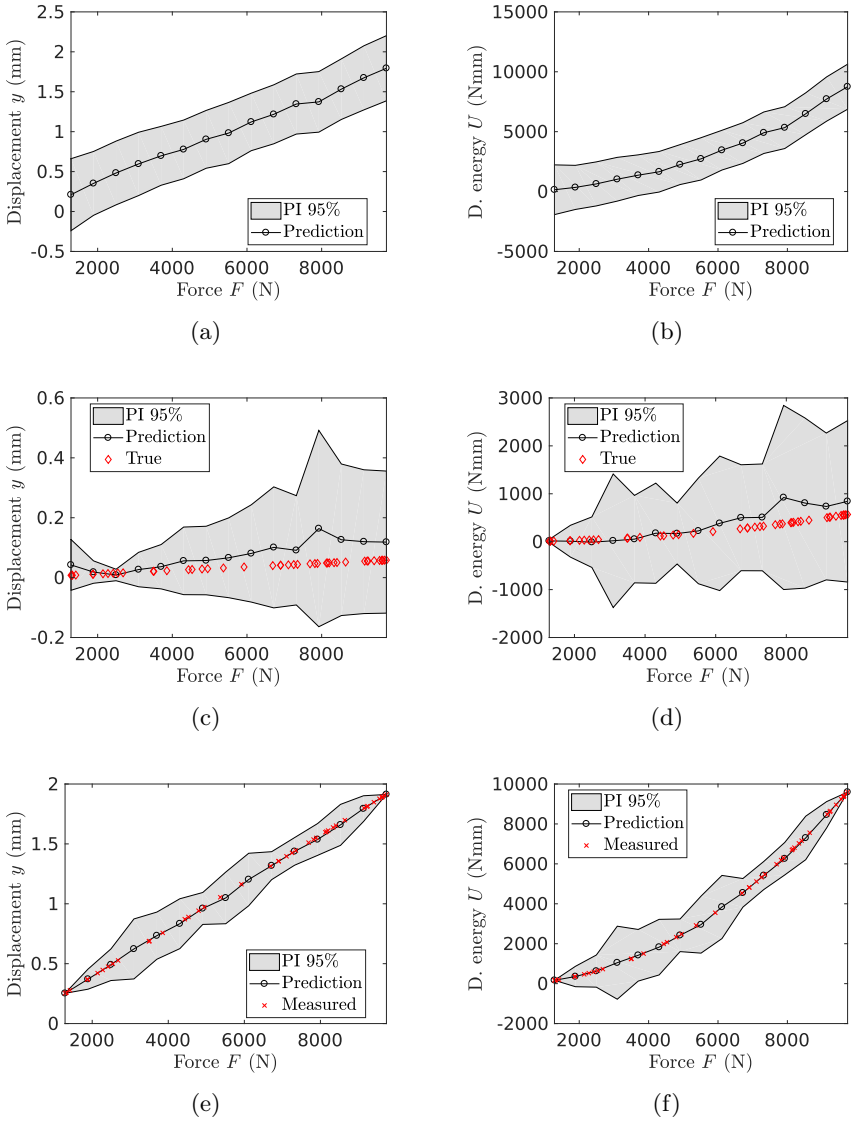
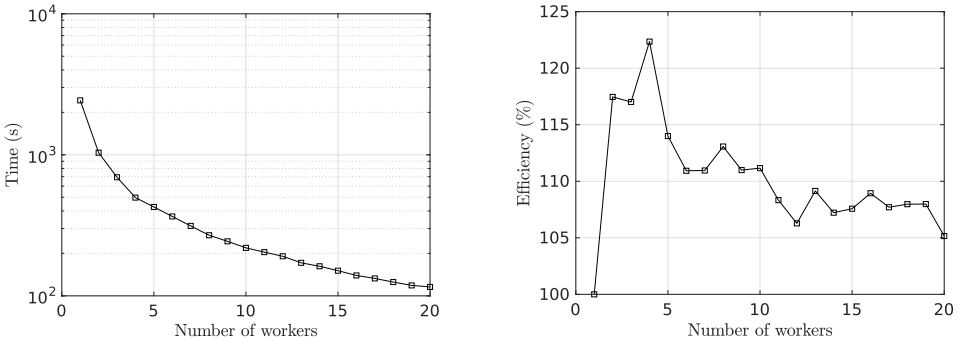


Figure 4.23: Predictions of the cantilever beam model (a) (b), discrepancy function (c) (d) and of experimental data (e) (f) for the displacement and deformation energy. True discrepancy function and experimental response represented as diamond markers.

Finally, and since in this example the numerical integration of Eq. (4.14) and (4.15) are computationally demanding, a comparison of computing time vs number of parallel processes is shown in Fig. 4.24(a). In particular, this information concerns the numerical integration of Eq. (4.17). The scaling efficiency of these computations (as a percentage of a linear scaling) can be calculated as

$$\text{Efficiency}(\%) = \frac{t_1}{N \times t_N} \times 100\% \tag{4.53}$$

where t_1, t_N and N represent the computing time for 1, N and the number of processing elements, respectively. Since in MATLAB a worker is defined as a session that performs the task computations, or worker process, these will be considered as the processing elements. Hence, the efficiency is plotted against the number of available MATLAB workers in Fig. 4.24(b). It can be seen that the efficiency grows up to four workers, starting to drop the linear scaling for a larger number. This usually occurs due to overhead communications and other algorithmic operations. The computing resources have been provided by the Cluster of Workstations operated by the Centre for Scientific Computing of the University of Warwick.



(a) (b)

Figure 4.24: Computational effort analysis for a different number of MATLAB workers used to integrate the correlation matrix in Eq. (4.17) for 170 simulated points and 70 observations. Computation time (a) and computational efficiency (b).

4.7.4 Observations with noise or residual variations

In this section, the capability of the modular Bayesian approach to address problems with noise is presented. Similarly to the previous sections an example of a cantilever beam with a point load applied to the beam tip is considered. The beam's deflection and deformation energy are simulated/observed in order to identify the Young's

modulus. Five signals with different levels of SNR, with -10, -3, 5, 11 and 18 dB, have been applied to the original noise-free signal. See Eq. (3.34) for a definition of the SNR. The original and polluted deformation energy process, for different levels of SNR, can be seen in Fig. 4.25.

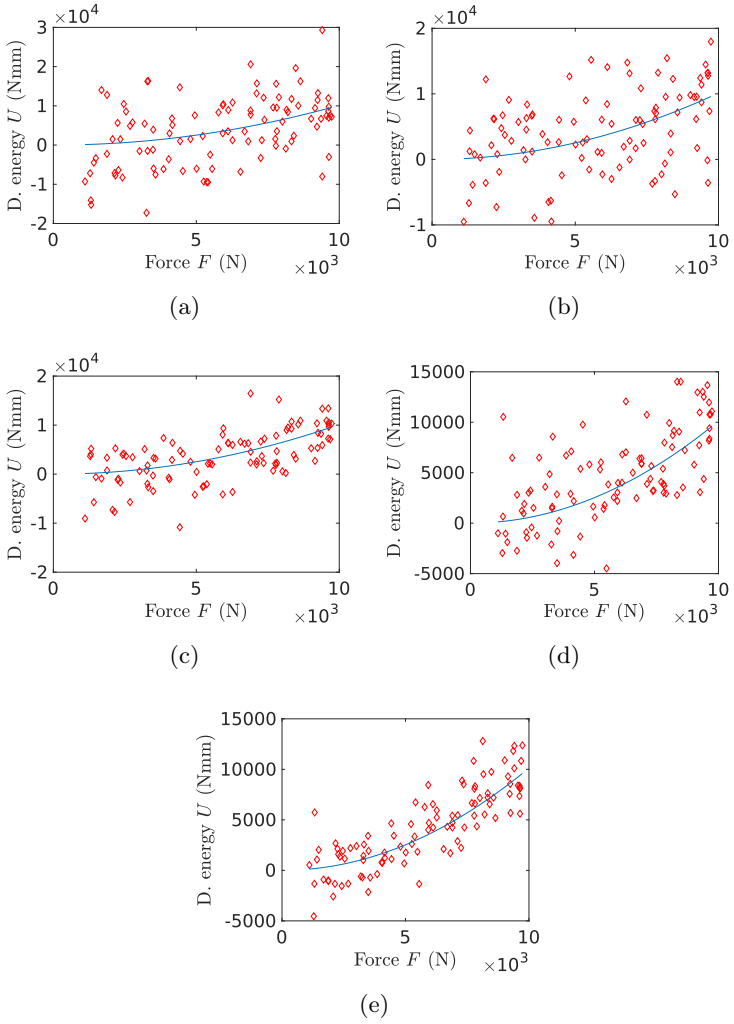


Figure 4.25: Observations (diamonds) of deformation energy for a SNR of -10, -3, 5, 11 and 18 dB in (a) (b) (c) (d) and (e), respectively, and unpolluted observations (solid line).

For this example, the trained mrGps will have their hyperparameters presented for each level of SNR. The number of observed and simulated data points used in the identification were 35 and 100, respectively. The cantilever model hy-

perparameters are similar to the ones presented in previous section. Namely

$$\hat{\boldsymbol{\omega}}_{F,M,E} = \begin{bmatrix} 0.295 & 0.195 \end{bmatrix} \; \hat{\boldsymbol{\beta}}_{y,U} = \begin{bmatrix} 0.134 & 0.341 \end{bmatrix} \quad \hat{\boldsymbol{\Sigma}}^2 = \begin{bmatrix} 0.276 & 0 \\ 0 & 4699646 \end{bmatrix}.$$

On the other hand, the signal with the most challenging SNR, i.e. -10 dB, was fitted with a discrepancy function mrGp with the following hyperparameters

$$\hat{\boldsymbol{\omega}}_{F,M} = 0.015 \quad \hat{\boldsymbol{\beta}}_{y,U} = \begin{bmatrix} -1.028 & -1.804 \end{bmatrix}$$

$$\hat{\boldsymbol{\Sigma}}^2 = \begin{bmatrix} 0.02 & 0 \\ 0 & 170160 \end{bmatrix} \quad \hat{\boldsymbol{\Lambda}} = \begin{bmatrix} 4.542 & 0 \\ 0 & 58132274 \end{bmatrix}$$

Since for this particular SNR, σ_{noise} = [4.214 70509653], the noise variance is in the correct range i.e. $\boldsymbol{\Lambda}$ = [4.542 58132274], although the deformation energy U was lower than the true value. To improve it, either the number of data points would have to be increased, or a lower value of the genetic algorithms fitness function tolerance would have to be set. The hyperparameters for the other SNR are shown in Table 4.5.

SNR (dB)	$\hat{\boldsymbol{\omega}}$		$\hat{\boldsymbol{\beta}}$		$\hat{\boldsymbol{\Sigma}}^2$		$\hat{\boldsymbol{\Lambda}}$	
	F	M	y	U	y	U	y	U
-3	0.009	0.009	0.596	-0.216	0.007	195434	2.005	32410936
5	0.484	0.484	0.266	-0.195	0.008	489668503	1.259	2986837
11	0.016	0.016	0.303	-0.298	0.005	92009	0.444	4223090
18	0.006	0.006	-0.189	-0.03	0.008	87752829	0.263	3527848

Table 4.5: Hyperparameters of discrepancy function mrGp for varying SNR.

The final posterior PDFs and its moments for each level of SNR can be seen in Fig. 4.26 and Table 4.6, respectively. As can be seen in the relative error column of Table 4.6, the worst performance registered, 38%, occurred at the lowest end of the scale -10 dB. In second place the 5 dB SNR registered an error of 5.16%. This error is below the remaining values, which are in the 6% range, and it occurred because of a convergence error of the genetic algorithms at a local minimum. The main conclusion is that there might be a more appropriate setting of the algorithm parameters in order to avoid convergence problems. These have not been explored in the present thesis.

Nonetheless, it is clear that under the proper circumstances, the methodology effectively accounts for observation errors. Moreover the present example includes model discrepancy at the same time, which makes is even more challenging. Finally,

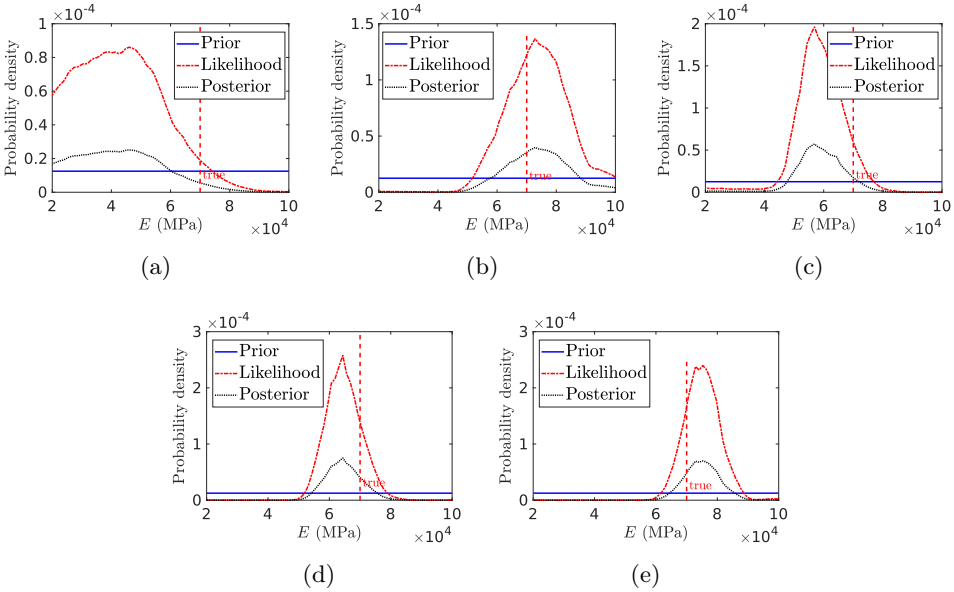


Figure 4.26: Inference of Young's modulus for different SNR. -10, -3, 5, 11, and 18dB in (a), (b), (c), (d) and (e), respectively.

SNR (dB)	E[θ] MPa	$\sigma[\theta]$ MPa	ϵ (%)
-10	43665	14231	37.62
-3	73618	10374	5.16
5	58781	5668.3	16.02
11	64721	5684.5	7.54
18	75045	5668.3	7.20

Table 4.6: Inference of Young's modulus for different SNR. Expected value, standard deviation and relative error.

it can also be noted that the standard deviation increases with a decreasing SNR. However, such an increase of estimation uncertainty can be improved by increasing the amount of observations, as shown in Fig. 4.27. In these figure, an increasing number of observations is supplied to the MBA, for a fixed SNR of 1dB. All the other input parameters are fixed.

4.7.5 Effects of large model discrepancy on identifiability

This section details the important topic of how a large model discrepancy can compromise the identifiability of the MBA. It is recalled that large model discrepancy is a frequent scenario in complex infrastructure.

In the cantilever beam example, this will be exemplified by considering the rotational spring at its supports with a lower stiffness. See Fig. 4.10 and Ta-

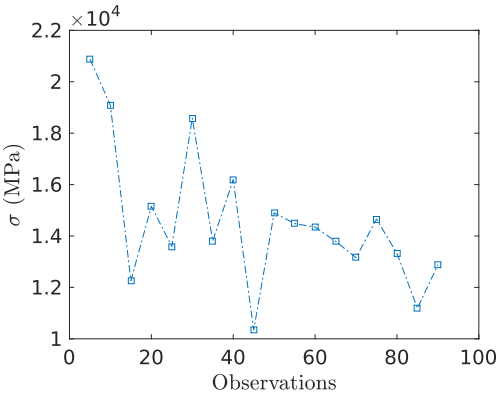


Figure 4.27: Relation between standard deviation of posterior distribution for an increasing number of observations with a SNR of 1dB.

ble 4.1 for clarification. The new stiffness will be ten times more flexible, i.e. $K^* = 15 \times 10^{10}$ Nmm/rad. The model mrGp was fit with 70 simulations and the number of observations has been changed parametrically, in six intervals from 5 to 75 observations, to assess the improvement of supplying more data.

Inference results can be seen in Fig. 4.28. The expected value of the Young's modulus has shifted towards a lower value indicating a more flexible material. In essence, the model is unable to separate the flexibility added by the spring, from the one attained because of the Young's modulus. Moreover, increasing the number of observations does not significantly improve identifiability, and the inference with the largest amount of data points had an expected value of $\mathrm{E}[\theta] = 54\,853$ MPa and variance of $V[\theta] = 5327^2$ MPa^2.

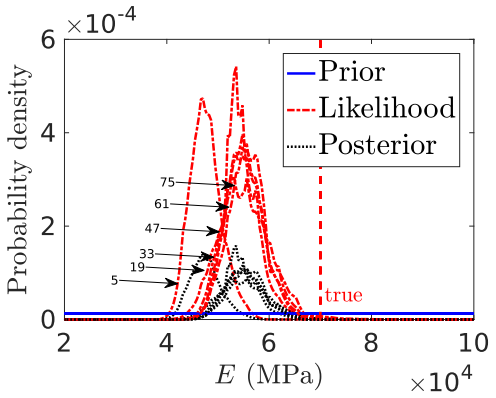


Figure 4.28: Inference of Young's modulus stiffness parameter in the presence of large model discrepancy. Arrows indicate number of observations of each inference curve.

Some improvement may be obtained, if for example, the end rotation of the beam is also considered. By doing so, the resulting posterior and likelihood PDF's are displayed in Fig. 4.29, and its moments are $\mathrm{E}[\theta] = 59\,106$ MPa and $V[\theta] = 10\,193^2$ MPa^2. Note that only 55 observations have been used to obtain these results, and the variance can be reduced by increasing this number. Finally, it becomes evident that the expected value has shifted towards the true value of the parameter, improving identifiability by $\varepsilon_r = \frac{59\,106 - 54\,853}{60\,000} = 7.1\%$.

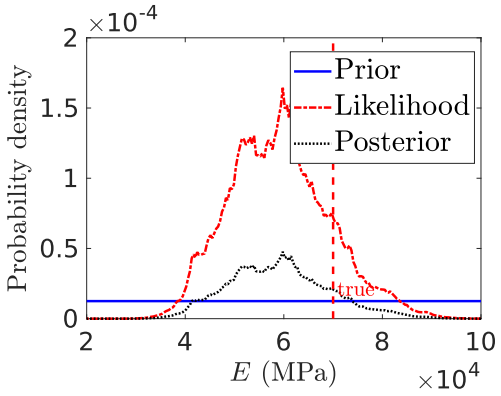


Figure 4.29: Inference of Young's modulus stiffness parameter considering the end rotation of the beam as an additional response.

The above example highlighted how the MBA can be affected by large model discrepancy, and how its identifiability can be improved by considering additional responses. As shall be seen in Chapter 6, the responses selection is not arbitrary, and through careful evaluation of the model it is possible to choose the responses which return the highest identifiability out of the MBA.

4.7.6 Effects of outliers

The previous examples have addressed the calibration of a model against data which follows a well behaved pattern. However it is very frequent to observe monitored data which does not follow such trends, and which are commonly referred to as outliers. Thus, in the context of this work, an outlier is interpreted as data that does not fit onto a regressive model, usually presenting a very large deviation relatively to the mean function of the model that is being regressed. In SHM, outliers can indicate diurnal shifts of the structural behaviour, malfunctions of the data-acquisition system, or in the worst case scenario, the presence of damage, e.g. a new nonlinear trend which is developing progressively from the regular structural behaviour.

Therefore, in the current section the cantilever beam model will be calibrated

against observations which contain outliers. The two signals of displacement and deformation energy presented in Section 4.7.4 have been contaminated with a SNR of 4.38 and 71.6 dB, respectively. Two outliers have been added as sample points within 6 standard deviations from the mean of the signals, i.e. $\mu + 6\sigma$, where μ and σ represent the mean and standard deviation of the two signals of interest.

To fit the cantilever beam model and its discrepancy function with mrGps, 40 and 100 training points have been generated, plus the two added outliers. The final hyperparameters of the mrGps are

$$\hat{\boldsymbol{\omega}}_{F,E} = \begin{bmatrix} 0.295 & 0.195 \end{bmatrix} \; \hat{\boldsymbol{\beta}}_{y,U} = \begin{bmatrix} 0.134 & 0.341 \end{bmatrix}$$

$$\hat{\boldsymbol{\Sigma}}^2 = \begin{bmatrix} 0.276 & 0 \\ 0 & 4700184.402 \end{bmatrix}$$

for the model, and

$$\hat{\boldsymbol{\omega}}_F = 0.014 \quad \hat{\boldsymbol{\beta}}_{y,U} = \begin{bmatrix} -0.666 & -0.219 \end{bmatrix}$$

$$\hat{\boldsymbol{\Sigma}}^2 = \begin{bmatrix} 0.004 & 0 \\ 0 & 291664.667 \end{bmatrix} \quad \hat{\boldsymbol{\Lambda}} = \begin{bmatrix} 2.67 & 0 \\ 0 & 9388641.902 \end{bmatrix}$$

for the discrepancy function mrGp.

The posterior PDF of the Young's modulus and the prediction of the processes are displayed in Fig. 4.30 and 4.31, respectively. Note that the outliers can only be observed in the predicted response plots, i.e. Fig. 4.31(c)4.31(d). Similarly to the previous examples it can be seen that the Young's modulus has been identified accurately, with a relative error of $\varepsilon_r = 1.56\%$. Although this error is larger than the one obtained in the first example, it should be noted that the current st-id is very challenging because of several sources of uncertainty. Namely one of the responses has a considerably low SNR, plus the added outliers, and the model discrepancy induced by the rotational spring.

Firstly, it can be seen that the algorithm is able to cope reasonably well with signals with different levels of observation error on each channel, where 95% of the observations are confined within the prediction intervals of the mrGp. Secondly, the two outliers can be seen clearly in Fig. 4.31(d), and as expected they are classed outside of the prediction interval. The main reason why the discrepancy function mrGp assigns very low probability density to the outliers, is precisely because they deviate grossly from the mrGp model, and its assumption of inherent smoothness of a process.

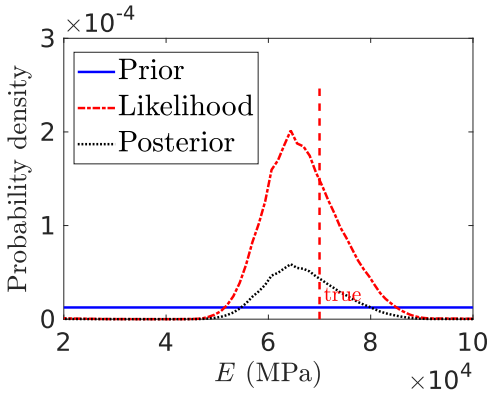


Figure 4.30: Inference of Young's modulus stiffness parameter considering three outliers in the observed data. Moments of the posterior are E[θ]=68903 MPa and σ[θ]=5495.9 MPa.

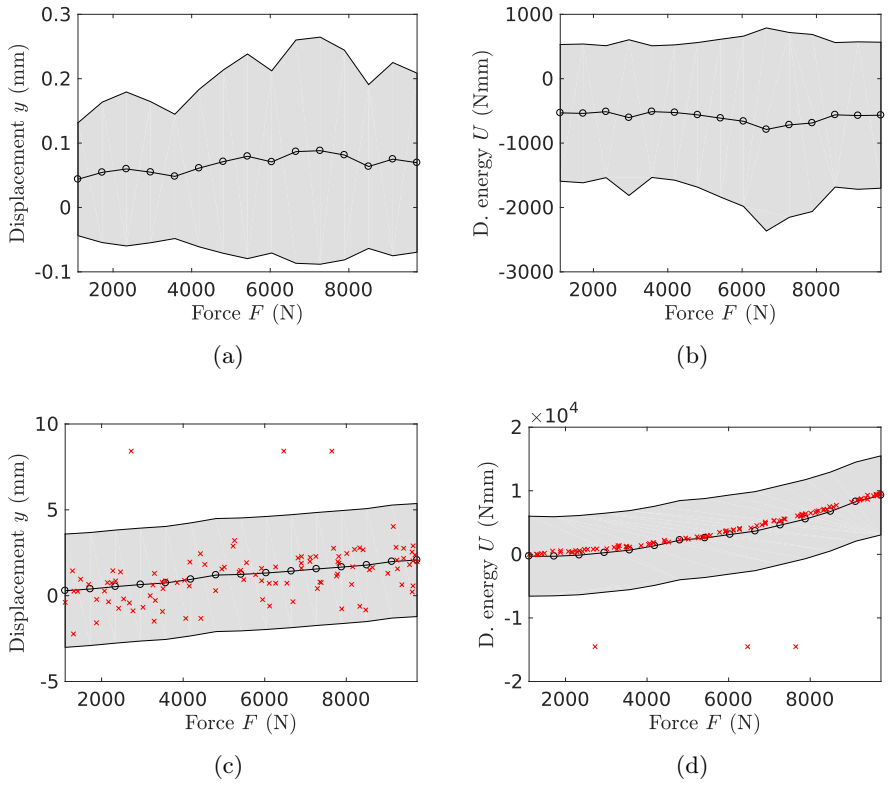


Figure 4.31: Predictions of the discrepancy function (top) and of experiment (bottom) for noisy signals, with outliers. Signals in (a) and (c) have a 4.38 dB and (b) and (d) 71.6 dB SNR. Observed data represented as crossed markers. Labels have been omitted in order to better visualise the outliers.

4.7.7 Identification using Gaussian prior information

It is recalled that the MBA formulation implemented in this work is able to numerically integrate the prior information onto the model mrGP, in module 2 of the algorithm, as described in Section 4.6.1. Such consideration allows the user to supply any prior PDF and to solve a multiple parameter calibration problem, in which it is necessary to perform multidimensional integration.

Since the examples in the previous sections have always used a uniform prior, in this section, identification of the cantilever beam example using a univariate and multivariate, i.e. a one dimensional and joint distribution, Gaussian prior is highlighted. The number of points and parameter values of these two examples are all described in Section 4.7.1 and Section 4.7.3. The considered priors will have the following properties

$$\mu_{ug} = 50{\times}10^3 \; \sigma_{ug} = 10{\times}10^3 \qquad \mu_{mg} = \begin{bmatrix} 53 \times 10^3 \\ 47.2 \times 10^7 \end{bmatrix} \; \Sigma^2_{mg} = \begin{bmatrix} 1 \times 10^8 & 0 \\ 0 & 4.5 \times 10^{15} \end{bmatrix} \tag{4.54}$$

where μ_{ug} and σ_{ug} are the mean and standard deviation of the univariate case in MPa units, while μ_{mg} and Σ^2_{mg} are the mean and covariance matrix for the multivariate case, in MPa and mm^4 units. Note that the criterion for setting the mean value/vector was to offset it relatively to the true value of the parameters, e.g. $E^* = 70\,000$ MPa for the Young's modulus, in order to make the identification more challenging. In short, we are providing an erroneous prior, which often happens in practice.

For the univariate case the inference of the Young's modulus is given in Fig. 4.32. It can be seen, that despite the inacurate prior information which has been supplied (solid blue curve), the likelihood function, i.e. the evidence, is able to correct the resulting posterior distribution. If more data is supplied, the likelihood weight would keep increasing over the prior and the inference would become more accurate. However, it should be noted that there is always a link between the hyperparameters of the discrepancy function (which builds the likelihood function) and the prior PDF, as explained in module 2 of the MBA. It can also be seen that the posterior is smoother, unlike the spiky likelihood, which occurs because of the natural regularisation of the posterior, when using an informative prior.

For the multivariate case, a similar situation occurs as seen in Fig. 4.33. Noticeably, as seen in the inertia inference plot, even if the prior information contains very inaccurate information the likelihood function PDF is near the true value.

Although the posterior PDF is not shown in this figure, its moments are

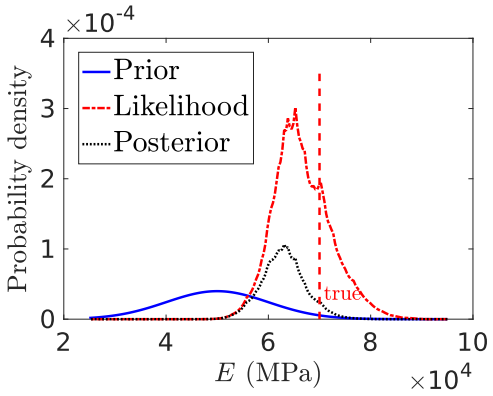


Figure 4.32: Inference of Young's modulus using a univariate Gaussian prior.

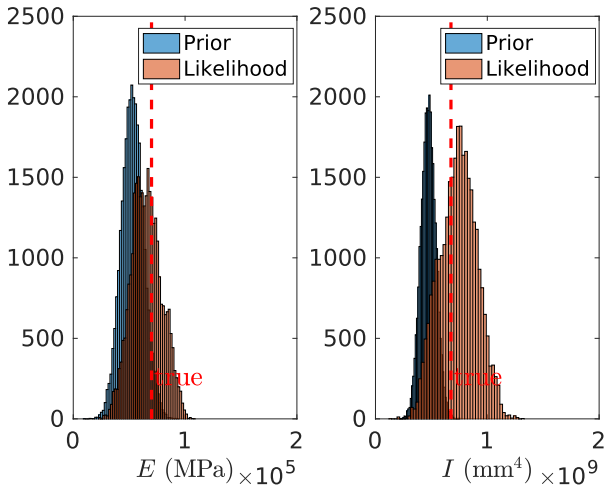


Figure 4.33: Inference of Young's modulus and cross-sectional inertia using a multivariate Gaussian prior – both PDFs have 30 thousand samples.

presented in Eq. (4.55). As can be seen these two quantities are negatively correlated, which is expected, since for a given deflection of the cantilever beam, an increase of rigidity by the Young's modulus parameter would imply a lower value of the cross sectional inertia parameter, to maintain the global bending stiffness EI constant.

$$\mathrm{E}[\theta] = \begin{bmatrix} 65517 \\ 7.4 \times 10^8 \end{bmatrix} \quad \Sigma^2[\theta] = \begin{bmatrix} 1.8 \times 10^8 & -1.2 \times 10^{12} \\ -1.2 \times 10^{12} & 2.7 \times 10^{16} \end{bmatrix} \begin{matrix} \text{MPa} \\ \text{mm}^4 \end{matrix} \tag{4.55}$$

These results suggest that the MBA algorithm is able to represent prior information in the form of a Gaussian PDF, performing Bayesian identification as expected. It should be noted that in its present form, the algorithm only allows for

uncorrelated Gaussian priors.

4.8 Summary

In this chapter, the core hybrid Bayesian framework of this thesis has been detailed. Developments over the original formulation, which allow to identify multiple parameters and an improved numerical performance, have also been described. It should be noted for example, that although Arendt et al.[146] considered an implementation of the MBA with multiple responses, their example only calibrated one parameter. These features, combined with the factorisation described in Section 3.7 constitute the novelty of the implemented approach.

Finally, an exhaustive set of numerical tests based on a simulated cantilever beam have validated the approach, and highlighted its different capabilities and drawbacks. This framework is now to be applied in real case-studies for both structural identification and measurement system design. These are the objectives of Chapter 5 and 6, respectively.

Chapter 5

Structural identification

5.1 Introduction

In this chapter the core methodology of this thesis, described in Chapter 4, is applied for st-id of a reduced-scale laboratory structure and a full-scale infrastructure.

Broadly speaking, the performance of st-id depends of the measured structural responses under scrutiny, how diversified they are, from which locations are they being measured, and how much data is monitored. Due to the spatial and temporal vastness that characterise a civil infrastructure, estimating an optimal sensor configuration is in itself a challenging endeavour. This is the subject of measurement system design, and shall be discussed in Chapter 6.

The performance of a probabilistic st-id also depends on how it handles uncertainties due to mechanical and geometric heterogeneity, non-stationary boundary properties, changing behaviour of moving systems, elevated intrinsic forces e.g., dead weight, uniqueness of the structural system, non-stationarity due to environmental and operational conditions, amongst others.

Given the above constraints, the current chapter details the first st-id practical applications of the enhanced MBA. Two specimens are examined thoroughly. A reduced scale aluminium bridge model subject to thermal loading (Section 5.2), and a full-scale infrastructure under traffic and temperature variations throughout a year period , the Tamar long suspension bridge (Section 5.3). The results include:

- identification of the aluminium bridge' supports stiffness and of the Tamar bridge main/side cables initial strain and stiffness of its bearings;

- prediction of strain/temperature responses for the aluminium bridge, which shall be presented in Section 5.2.3;

- prediction of discrepancy functions of the Tamar bridge natural frequencies and its mid-span displacements, in Section 5.4;

The aforementioned results will be compared against reference values or available monitored data, in order to provide further insight of the performance of the MBA. To the best of the author's knowledge, these tests are the first practical applications of the enhanced MBA, particularly for temperature-based structural identification. A short summary of the key conclusions of this chapter shall be presented in Section 5.5.

5.2 Aluminium bridge subjected to thermal loading

In this case-study a reduced-scale laboratory aluminium bridge inspired by the New Joban Line Arakawa (Japan) railway bridge, was built at the Warwick Civil Engineering Laboratory and subjected to thermal loading due to infrared heaters. Some of the advantages of using the reduced scale model over the simulated examples of Section 4.7 or the full-scale infrastructure are:

1. more realistic conditions, e.g. noise, inherent randomness and residual deformation of temperature loading, comparatively to the simulated examples;

2. known structural parameters can be used to test the reliability of the methodology and validate the st-id results; possibility of easily testing different measurement scenarios, comparatively to the Tamar bridge;

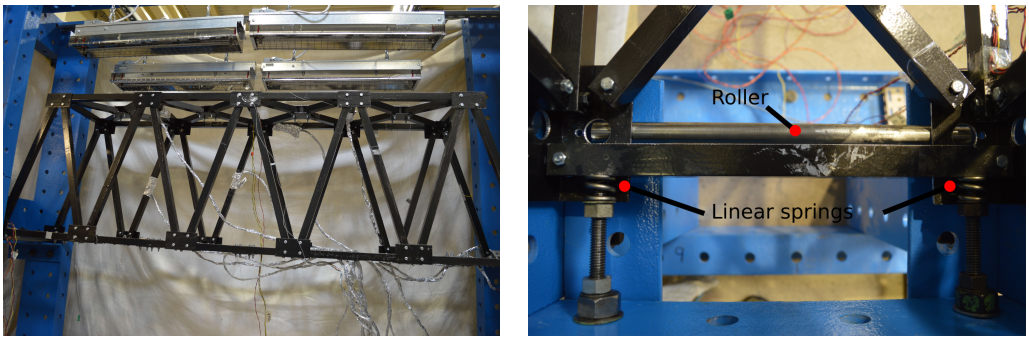
3. damaging the reduced-scale structure is permissible and allows to easily test damage identification methodologies, which is generally not allowed in full-scale infrastructure.

Typical daily ambient temperature in the laboratory ranged from 18°C up to 21°C. A numerical model of the structure was also developed to study the phenomena. The stiffness of a pair of springs located at one of the ends of the bridge will be considered as a model parameter to be calibrated. Subsection 5.2.3 details a combinatorial analysis to select the best out of a set of available inputs, to maximise the performance of the method, which is subsequently applied and results are shown in subsection 5.2.4.

5.2.1 Experiment

The truss structure in Fig. 5.1(a) is simply supported at its ends and is constrained on one of them by two linear springs Fig. 5.1(b). Its geometrical properties will be

detailed along with its monitoring system.



(a) (b)

Figure 5.1: Aluminium bridge subjected to thermal loading (a) and detail - springs constraint (top view) (b).

Beforehand, some information about the restraining springs will be detailed. Namely, the supplier is Lee Spring and its properties are discriminated in Table 5.1. Additionally, the tolerances of the spring rate are ±10%, i.e. a confidence interval of [497.03 607.49] N/mm.

Part Number	LHP 192L 01S	-
Outside Diameter	24.637	mm
Hole Diameter	25.400	mm
Wire Diameter	4.876	mm
Load At Solid Length	1742.726	N
Free Length	19.049	mm
Rate	552.26	N/mm
Solid Length	15.900	mm
Rod Diameter	12.700	mm
Number of Coils	1.2	-
Total Coils	3.2	-
Finish	PASSIVATE PER ASTM A967	-
Material	17-7	-

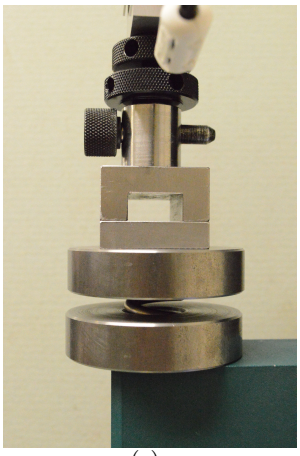
Table 5.1: Properties of tested spring (according to Lee Spring).

In addition to the above mentioned information, one of the springs has been tested in a compression Tinius Olsen model 25ST testing machine (cf. Fig. 5.2(a)), equipped with a ± 25 kN load cell. In order to account for the precision of the machine, eight tests of compressive loading have been performed up to a maximum extension of 5 mm at a speed rate of 2mm/min. Finally, in order to account for the irregular shape of the spring, at each test the spring restraints (cf. Fig. 5.2(b)5.2(c))

have been rotated over the loading axis.

(a)

(b)



(c)

Figure 5.2: Test set up: Tinius Olsen 25ST materials testing machine (a), spring restraints (b) and final restrained spring ready for testing (c).

The loading curves of the eight tests are plotted in Fig. 5.3 (crossed dots). One of the eight curves is offset relatively to the others, because of the irregular shape of the spring and its consecutive rotation over the loading axis. Note that the initial phase of the loading curves (position below 1 mm) has been removed, to avoid initial nonlinear trends in which the spring force is still being mobilised. Finally, a linear regression model defined as

$$F = 500.43 \times d - 342.02 \tag{5.1}$$

where F is the load acting on the spring and d is its vertical displacement, has been fitted to the data and plotted along with it (solid line). The slope of this model indicates that the spring stiffness has a mean value of 500.43 kN/mm, which is within the above tolerance.

Geometrical information and the measurement setup of the bridge are displayed in Fig. 5.4. The strain gauges have been installed before the bridge components were assembled, and therefore the measurable strain includes the self-weight of

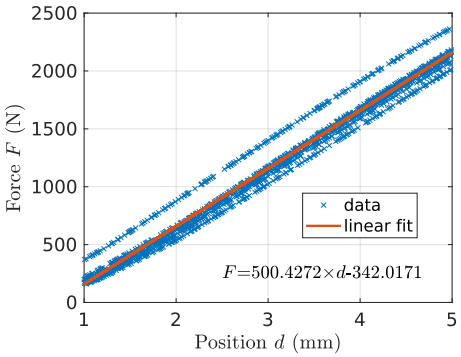


Figure 5.3: Linear regression of compressive spring load test.

the whole structure plus the effects of thermal expansion and possible connection influence. Material properties of the bridge are: aluminium alloy of grading 1050AH14 and 6082T6 for the gusset plates and for the box section beams, respectively. Stainless steel M6 8.8 bolts connect the multiple members of the bridge. Measurements with eleven strain gauges and four thermocouples took place during the experiment at a sampling rate of 1 Hz. A proportional-integral-derivative (PID) controller on a Labview routine was used to control the infrared heaters.

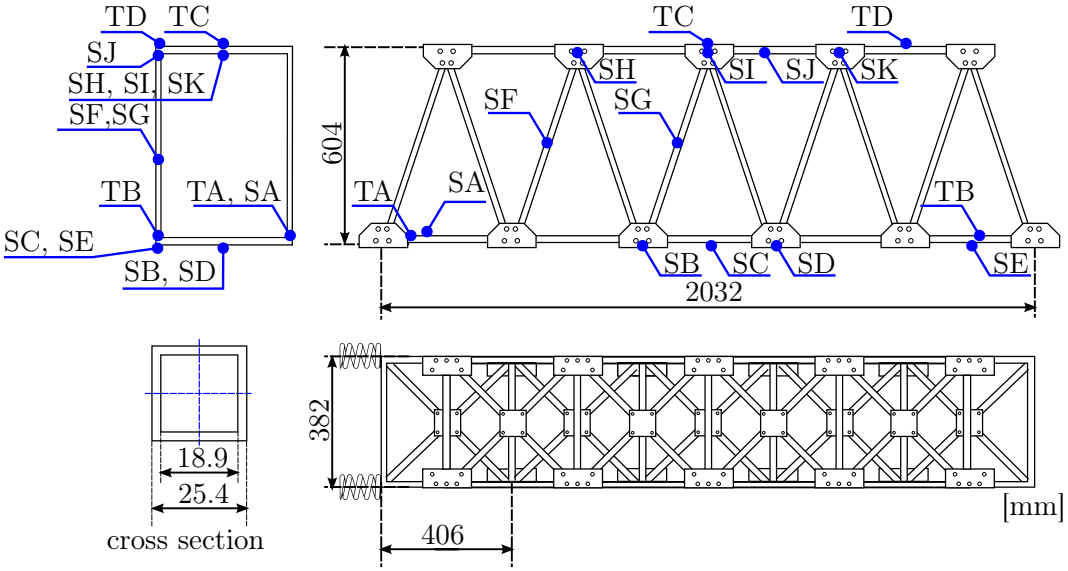


Figure 5.4: Aluminium bridge measurements diagram (dimensions in mm) - thermocouples TA-TD and strain gauges SA-SK. Placement of labels above/below bar reflects the sensor position.

Temperature and strain readings during the main experiment, which took approximately half an hour, are displayed in Figs. 5.5(a) and 5.5(b). The reason

why the strain even at the top is in compression, is that all the strain gauges have been placed on the bottom side of the bars. Therefore, despite the global bending of the structure that leaves the top bars under tension there is a localised bending at the top bars, which is measured as a compression on their bottom side.

The temperature-strain relation visible on Fig. 5.5(b) is not linear, because the measurements are performed on the surface of a squared hollow section, which cools down faster than the internal cross section and will, therefore, still have some residual thermal deformation when cooling down.

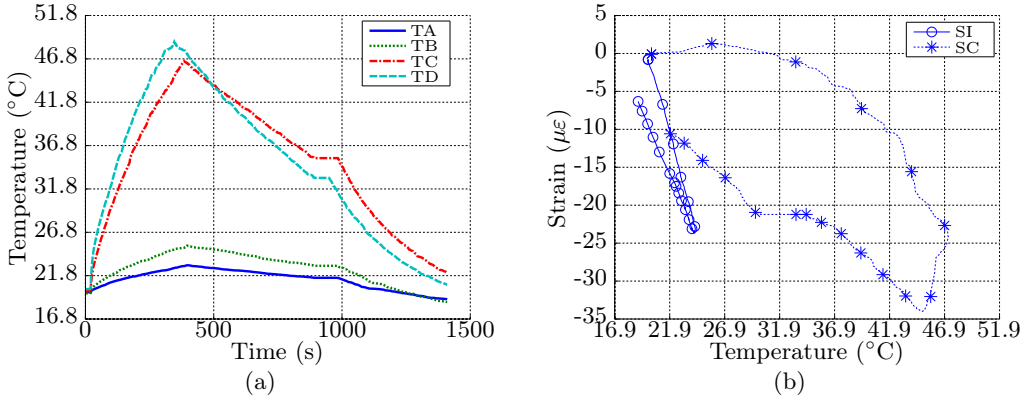


Figure 5.5: Aluminium bridge heating/cooling cycle - Temperature readings (a) and strain measurements (b), see Fig. 5.4 for reference.

Although the strain gauges are thermally compensated for aluminium, the data from the strain gauges has been post-processed to remove the remnant thermal output effect, which is related to the natural thermal expansion of the gauge.

5.2.2 Finite element model of aluminium bridge

The model was developed in ANSYS and coded using APDL (ANSYS parametric design language) [203]. Beam elements with rotational stiffness were used to represent the bars of the bridge. The material model is isotropic, linear-elastic with Young's modulus $E = 70 \times 10^6$ kPa, Poisson coefficient $\nu = 0.35$ and coefficient of thermal expansion $\alpha = 23.1 \times 10^{-6}$ °C^{-1}, as standard aluminium. Reference temperature was set as $T_0 = 20$°C. A uniform distribution of temperature was applied through all the bar elements, and is based on the mean of the thermocouple measurements at the top of the bridge seen in Figure 5.5(a). Essentially, nine linearly spaced points of thermal loading from 20.19°C to 45.49°C were simulated as a quasi-static analysis, with different values of the spring stiffness. Each analysis took approximately 0.128

seconds on an Intel i7 quadcore 2.2 GHz, 6 GB of RAM and an SSD drive.

Figure 5.6 shows the strain output of the bridge model for this loading condition. It is considerably easier to model the infrastructure behaviour with a uniform temperature gradient on all of its elements, but obviously this is a model discrepancy, since in the laboratory experiment the top of the bridge is much hotter than the bottom side, as seen in Figure 5.5(a). Since the strain of the FE model is being sampled at the bottom side of the top bars, the effect of the localised bending should be relatively small on the results.

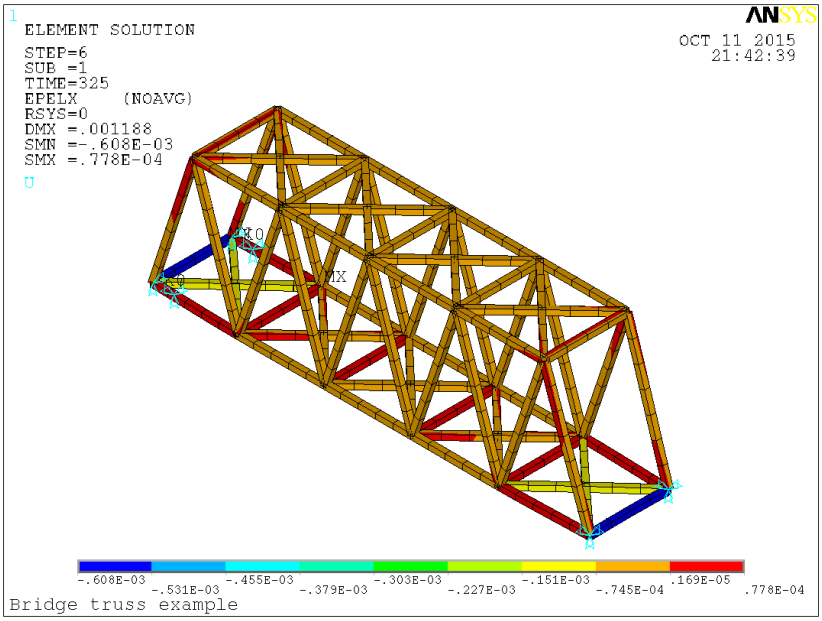


Figure 5.6: FE model with linear springs at support, maximum temperature of heating cycle.

Finally, input data used for st-id is shown in Table 5.2 and the output are the moments from the posterior distribution of the spring stiffness. This will be shown in Section 5.2.4.

5.2.3 Sensor location combinatorial analysis

This section presents a study of the influence of sensor location on the capability of inferring the spring stiffness true value. The present analysis is illustrative of which responses provide better identifiability, and shall be more thoroughly discussed in Chapter 6.

The MBA was applied to infer the spring stiffness, based on all possible combinations of two out of eleven strain measurements available from the laboratory

Parameter	Description
$\boldsymbol{X}^e$	13 readings of temperature at the top of the bridge (mean of TC and TD, Fig. 5.5(a)) from the beginning of the heating cycle until its maximum
$\boldsymbol{Y}^e$	strain measurements at points B and G (Fig. 5.4) for each of the temperature readings
$\boldsymbol{\theta}$	spring stiffness with $\theta^* = 552.26$ N/mm
$[\boldsymbol{X}^m, \boldsymbol{\Theta}]$	combination set of each temperature reading with possible spring stiffness values on a 9×9 input grid space with K[300; 1000] N/mm and T[20.19 to 45.49]°C
$\boldsymbol{Y}^m$	simulations from the numerical model for $[\boldsymbol{X}^m, \boldsymbol{\Theta}]$

Table 5.2: Dataset for aluminium bridge.

experiment. Only the first three modules of the algorithm are necessary for this combinatorial analysis. The resulting posterior distribution is a Gaussian-shaped distribution whose moments (mean value and standard deviation) estimate the stiffness value. Therefore there are $C_2^{11} = 55$ possible combinations of responses, and each took approximately 16.50 s on a laptop computer. From the whole set the combinations with minimum standard deviation $\sigma[\theta]$ and expected value $\mathrm{E}[\theta]$ closer to the spring stiffness real value were selected. Such selection is possible only because the true value of the spring stiffness is known, and ϵ(in %) can be determined.

Table 5.3 displays the results of the combinatorial analysis with a relative error $\epsilon = \frac{|K - \mathrm{E}[\theta]|}{K} \leq 10$ %. The successful convergence by change in the fitness value (Yes or No in the right column) of the genetic algorithms (GA) for the maximum likelihood optimisation of module 2 is also shown.

S1	S2	$\mathrm{E}[\theta]$	ϵ(%)	$\sigma[\theta]$	GA	S1	S2	$\mathrm{E}[\theta]$	ϵ(%)	$\sigma[\theta]$	GA
A	H	605.04	9.56	171.87	Y	D	H	520.64	5.72	147.19	Y
A	I	596.49	8.01	169.60	Y	H	I	603.12	9.21	169.62	Y
A	J	602.45	9.09	171.56	N	H	J	599.60	8.57	170.28	Y
A	K	603.98	9.37	170.29	Y	H	K	603.99	9.37	161.07	Y
B	G	532.03	3.66	47.72	Y	I	J	598.17	8.31	169.46	Y
C	E	547.31	0.90	102.21	Y	I	K	604.15	9.40	168.44	Y
D	G	533.90	3.33	126.47	Y	J	K	599.37	8.53	169.94	Y

Table 5.3: Seven best two strain measurement combinations with $\epsilon \leq 10\%$ for inference of the spring stiffness using the MBA.

On the other hand Table 5.4 shows the results obtained for standard deviations $\sigma \leq 70$ N/mm. Finally the worst values of metrics for σ and ϵ were 171.87 N/mm and 43.18 %, respectively. The results show that:

- the combination of responses measured at point B and G provide the lowest error and variance on the inference of the spring stiffness. Although the combination C and E has a smaller error its variance is above the maximum allowed of $\sigma \leq 70$ N/mm, having henceforth been discarded;
- typically combinations involving measurements near bridge ends (A and E) either present a high variance (notice that no combination with E is present in Table 5.4), or the genetic algorithm optimisations do not converge (see the N on the GA column of Table 5.4). This indicates that measurements near model singularities such as supports tend to present poor identifiability;
- in general combining one measurement from the bridge bottom with one of the middle or top gives the lowest error and variances, which indicates that combining locations with a different loading range enhances identifiability.

S1	S2	E[θ]	ϵ(%)	σ[θ]	GA	S1	S2	E[θ]	ϵ(%)	σ[θ]	GA
A	C	430.84	21.99	58.79	N	D	F	433.96	21.42	65.98	Y
A	F	444.54	19.51	68.20	N	F	G	640.50	15.98	67.50	Y
A	G	687.24	24.44	47.97	N	F	I	439.67	20.39	65.44	Y
B	D	394.75	28.52	68.56	Y	G	H	681.50	23.40	52.91	Y
B	G	532.03	3.66	47.72	Y	G	I	684.32	23.91	50.80	Y
C	G	608.89	10.25	58.89	Y	G	J	676.23	22.45	55.87	Y
C	I	437.85	20.72	56.95	Y						

Table 5.4: Seven best two strain measurement combinations with $\sigma \leq 70$ N/mm for inference of the spring stiffness using the MBA.

To further support this interpretation, a comparison between an identifiability metric applied by Arendt [146] in his simulation, against this example shall be carried out in the following section. A more thorough analysis of the best measurement system design, which does not require prior knowledge of the property to be identified or the measured responses, will also be carried out in Chapter 6.

5.2.4 Model calibration results and discussion

On the basis of the results obtained above, in this section a detailed calibration of the FE model against the laboratory scale aluminium bridge data, will be presented, using the the strain at locations B and G. The design variable, response outputs and structural parameter are equivalent to what was presented in Table 5.2. The polynomial regression functions are set as a constant value $\boldsymbol{H}(\bullet) = 1$, as detailed in Table 3.1, and the prior of θ as a uniform probability density function (PDF).

A mrGp of the model is presented in Figs. 5.7(a) and 5.7(b), for the response surface of the strain at locations B and G, and its hyperparameters are shown below. Notice how the uncertainty cloud of the prediction interval shrinks and increases relatively to the distance to the input dataset, to account for the uncertainty of the numerical model.

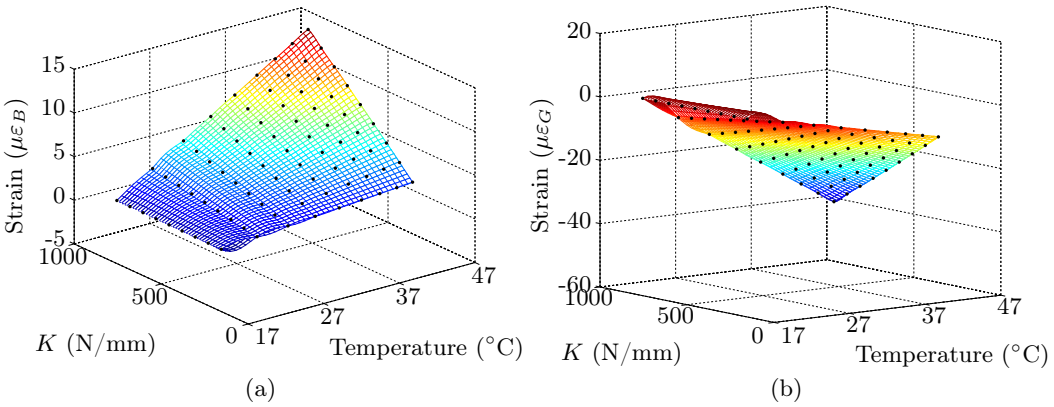


Figure 5.7: 95 % prediction interval of the posterior multiple response Gaussian process - 40×40 grid with 9×9 grid input set (black dots) of strain at position B (a) and position G (b).

The following data represents estimates of parameters, that fully characterise the Gaussian processes, approximated on the modular Bayesian approach. They are named here as hyperparameters $\boldsymbol{\phi}$ and comprise a matrix of regression coefficients $\boldsymbol{\beta}$, a variance matrix $\boldsymbol{\Sigma}$, a noise variance matrix $\boldsymbol{\Lambda}$ and the roughness parameters $\boldsymbol{\omega}$. The numerical model mrGp has hyperparameters

$$\hat{\boldsymbol{\omega}}_{T,K} = [2.0 \;\; 1.4] \qquad \hat{\boldsymbol{\Sigma}} = \begin{bmatrix} 3.0645 & 8.4687 \\ 8.4687 & 25.3702 \end{bmatrix}$$
$$\hat{\boldsymbol{\beta}}_{SB,SG} = [-0.105 \;\; 0.121]$$

and the discrepancy function mrGp has hyperparameters

$$\hat{\boldsymbol{\omega}}_T = 6.87 \qquad \hat{\boldsymbol{\Sigma}} = \begin{bmatrix} 22.90 & 1.07 \\ 1.07 & 2.75 \end{bmatrix} \qquad \hat{\boldsymbol{\Lambda}} = \begin{bmatrix} 0.27 & 0.03 \\ 0.03 & 0.59 \end{bmatrix}$$
$$\hat{\boldsymbol{\beta}}_{SB,SG} = [-3.766 \;\; -0.242]$$

The result of the identification task (i.e. inference of the spring stiffness) is shown in Figure 5.8. A deterministic model identification approach (Det. label in the legend), based on the minimum value of the sum of root mean squared errors (MSE) of the two strains (positions B and G), has a minimum stiffness value at $K = 661.5$ N/mm with an error of 15.16 $\mu\varepsilon$. This error is more than five times

larger than that of the MBA (2.82 $\mu\varepsilon$). Furthermore as seen in Figure 5.8(b), the stiffness value resulting from the MBA, 532.03 N/mm, is closer to the real value of stiffness given by the supplier 552.26 N/mm (3.66%) and the value estimated from testing 500.43 N/mm (6.29%) compared to the stiffness identified using the deterministic model identification.

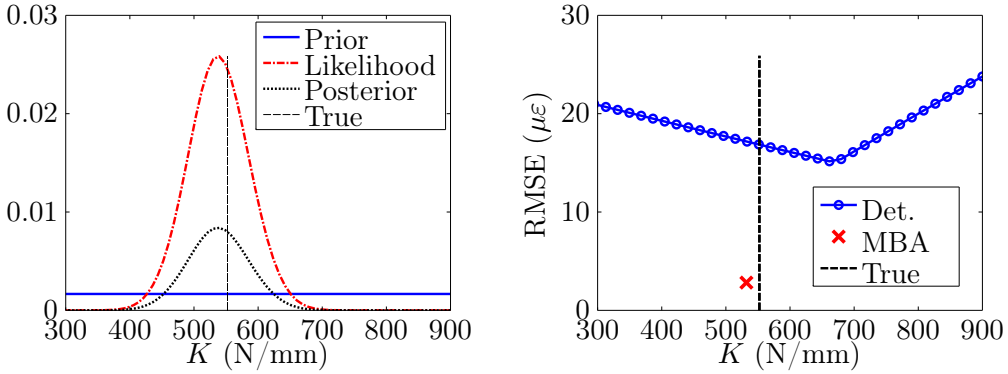


Figure 5.8: Inference of the stiffness of linear springs at aluminium bridge end, against reference value (vertical dashed line) for MBA (a) and comparison of deterministic approach and MBA by root mean squared error (b).

Note also that there exists some additional friction at the bridge supports, between the gusset plates and the installed roller (cf. Fig. 5.1(b)) which can easily explain why the identified value is larger than the one obtained with the linear model in Eq. (5.1). These results demonstrate the superior performance of MBA not only in terms of identifiability but also in the ability to predict structural responses. Despite the fact that the responses used for identification have the same characteristic nature (i.e. strain) the mean value and the variance closely approximate the real stiffness value. It is expected that if other responses such as displacements were given as input this uncertainty would be further improved.

By comparison of these results against the hierarchical Bayesian framework from Behmanesh [124], it can be seen that the current estimate only deviated by 3.66 % or 6.29%, whereas the hierarchical Bayes estimated model parameters with almost 20 % deviation from the true values (see Table 3 of [124] for clarification). Also to further justify that the improvement on identifiability was achieved, mainly due to the sensor position and not to the nature of the measured responses, a comparison between Arendt's metric in his simulated example and the current analysis was applied here. Essentially, an improvement of identifiability by using multiple responses should be quantifiable through the posterior standard deviation of the model

parameter that is being calibrated, following the formula $(\sigma^{\mathrm{SR}}_{i,j,\mathrm{min}} - \sigma^{\mathrm{MR}}_{i,j,\mathrm{post}})/\sigma^{\mathrm{SR}}_{i,j,\mathrm{min}}$, where $\sigma^{\mathrm{SR}}_{ij,\mathrm{min}} = \min(\sigma_{i,\mathrm{post}}, \sigma_{j,\mathrm{post}})$ is the minimum posterior standard deviation of the calibrated parameter with individual responses, and $\sigma^{\mathrm{MR}}_{i,j,\mathrm{post}}$ is the posterior standard deviation with multiple responses. In the current example an improvement of 10.1 % was observed, which is similar to that of 14 % obtained by Arendt (see Table 3 of [146]) on his simply supported beam for responses with similar nature.

Based on the inference of the spring stiffness the mrGp metamodel predictions for strain B and G are shown in Figure 5.9 and 5.10, respectively. Notice how the strain at B has a positive trend for the model prediction opposite to the negative slope in the experimental prediction, and how the considerable discrepancy was predicted accurately on Figure 5.9(b). The solid lines with diamond markers represent the measured response, which is within the uncertainty region. Since this region accounts for the uncertainty of the spring stiffness, model discrepancy, noise and the fact that the model response is only known at a set of discrete points (code uncertainty), the true undisturbed process should also fall within that limit with a 95 % accuracy.

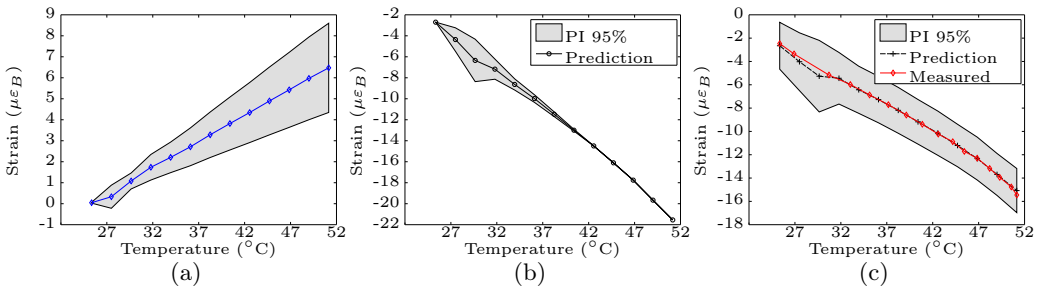


Figure 5.9: Strain at position B - Prediction interval 95 % confidence for numerical model 5.9(a), discrepancy function 5.9(b) and experimental response 5.9(c).

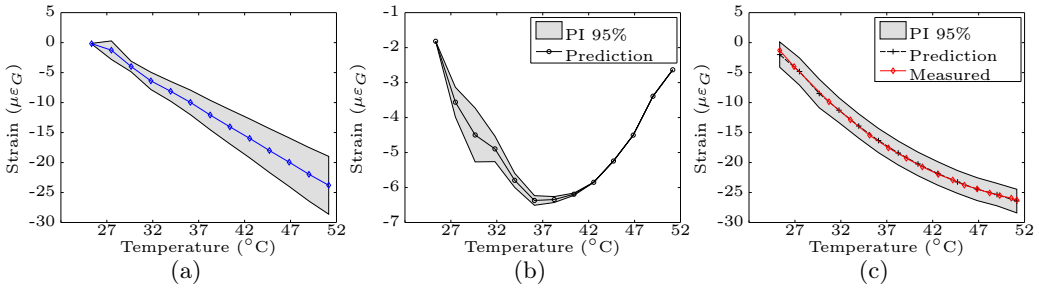


Figure 5.10: Strain at G aluminium bridge - Prediction interval 95 % confidence for numerical model 5.10(a), discrepancy function 5.10(b) and experimental response 5.10(c).

This section concludes the study of the application of the MBA for st-id of the aluminium truss bridge example. The structural stiffness of the supports of the example have been accurately identified, based on the two most informative responses of strain available. In the following section a much more complex identification problem of the Tamar suspension bridge shall be detailed.

5.3 Tamar bridge experimental and simulated dataset

In this section a summary description of: the Tamar bridge's SHM system; experimental data under consideration for st-id; the FE model developed to study its behaviour; and relevant assumptions/simplifications are detailed.

Tamar bridge is a 335 m long suspension bridge, built in 1959, and reconstructed between the 1999-2001, where a larger side deck and stay-cables have been added, see Fig. 5.11 for a reference. Two long-term monitoring systems and several expeditions have been carried out through time for reliability assessment. Additionally, an FE model was developed by Westgate [204] to study environmental and operational effects on its structural performance. It is worth mentioning the following excerpt "*In fact, Tamar Bridge has so far presented a challenging case study for model calibration, lending support to the view that no single model provides a perfect representation of a structure when matched to provided experimental data*" [205].

(a) (b)

Figure 5.11: Tamar suspension bridge, before 1978 [206] (a) and after 2012 [207] (b) its reconstruction in the late 1990s. Note the larger deck and stay cables which have been added in 2001.

On the basis of the methodology described in Chapter 4 and the aforementioned data/model, three key properties, which are deemed relevant for a stake holder of the infrastructure, will be estimated. One is the friction in the thermal expan-

sion bearings of Saltash tower, which can lead to deck cracks and further structural anomalies. The remaining two properties are the initial strain in the main and stay cables of the bridge. The initial strain is a critical property for a suspension bridge, since its increase could indicate internal damage, such as broken wires, corrosion, cracks and wear, and is defined as the strain relative to when the bridge cables have been installed initially, i.e. containing all the load-history that the cables have supported since installation.

5.3.1 Monitored data and post-processing

It is important to establish which design variables are relevant to study Tamar bridge's dynamic behaviour. A study from Cross et al. [208] indicates that traffic, temperature and wind have the most influence on Tamar bridge natural frequencies, by decreasing order of relevance. However, the present study will be limited to the effects of traffic and temperature.

In the absence of the above information, a normal procedure would include:

- monitor the structural behaviour for a certain period of time, preferably at least for a year period;
- analyse existent correlations between environmental/operational effects and the structural output, displacements, vibration data, etc.;
- select the effects which have the highest influence on the structural output and consider them during the modelling process.

Several sensors have been installed through time on Tamar bridge, but for the purpose of the present study, it suffices to consider data which was monitored from a set of accelerometers, a total positioning system (TPS) reflector, and thermocouples which are visible in Fig. 5.12. The available data also includes vehicle counts from toll gates of the Plymouth side. The monitoring period ranged from May of 2009 to March of 2010, where synchronised temperature/traffic/modal data was found to be richer. Furthermore, a year time-frame was assumed as a good reference for calibration/validation of the FE model. Relevant post-processing operations will now be detailed.

Firstly, the natural frequencies of the structure were determined with a Stochastic Subspace Identification (SSI) technique [209], based on available acceleration data. Specifically, at each half-hour, ten minute acceleration recordings were post-processed to determine the natural frequencies and mode shapes of the infrastructure. See Section 4 of [208] for a more extensive description of the procedure.

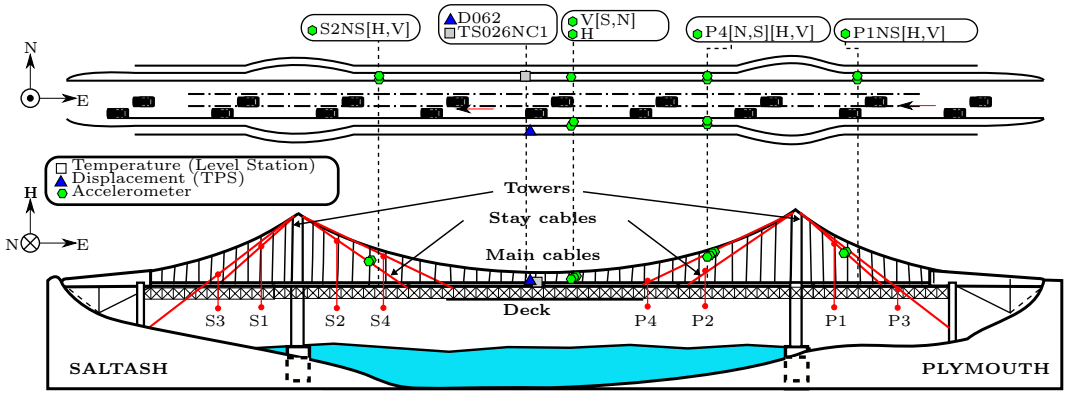


Figure 5.12: Diagram of Tamar bridge SHM system – cable temperature sensor, displacement reflector and accelerometers from whose natural frequencies/mode shapes are estimated. There are 16 stay cables on North/South and Saltash/Plymouth sides.

The Bayesian framework from Zhu and Au [210] would be an interesting alternative to obtain such modal properties. Secondly, displacements at the middle span of the bridge in three directions, vertical, East and North, were obtained from the TPS. Finally, sampling has been synchronised for the whole dataset, and despite the MBA ability to correctly identify outliers these were also removed. For simplicity, it has been assumed that the outliers have been caused by malfunctions of the data-acquisition system. Speculatively speaking, considering the outliers in the data set would make little difference in the final result, since they would have been assigned a low probability density by the mrGp, as shown in the numerical example of Section 4.7.6. However care should be exercised, since a structural anomaly could be the cause of the outliers, and an important indicator of the true structural behaviour could be overlooked. An alternative way to consider the outliers presence would be to combine the MBA with a different type of Gaussian process, such as the Bayesian regression trees model from Worden and Cross [151].

After application of the operations mentioned above to the original raw data, a 2419 points dataset was obtained, which can be visualised in Fig. 5.13. Visible trends indicate linear correlations, except for the traffic/displacement relation in Fig. 5.13(d). Therefore, a linear correlation function (or kernel) was assumed for the mrGps that fit the discrepancy function and the FE model [184]. Furthermore, the zero displacements at highest temperatures in Figs. 5.13(c) and 5.13(d) occur because the data has been offset relative to the highest peak of temperature $\times$ traffic load. Frequency labels follow the convention: 'L' is a lateral mode shape, 'V' is vertical mode shape, 'T' is a torsional mode shape, 'TRANS' is a longitudinal translation

mode, 'S' is symmetric, 'A' is asymmetric, 'SS' is side span and the numbers are their relevant order.

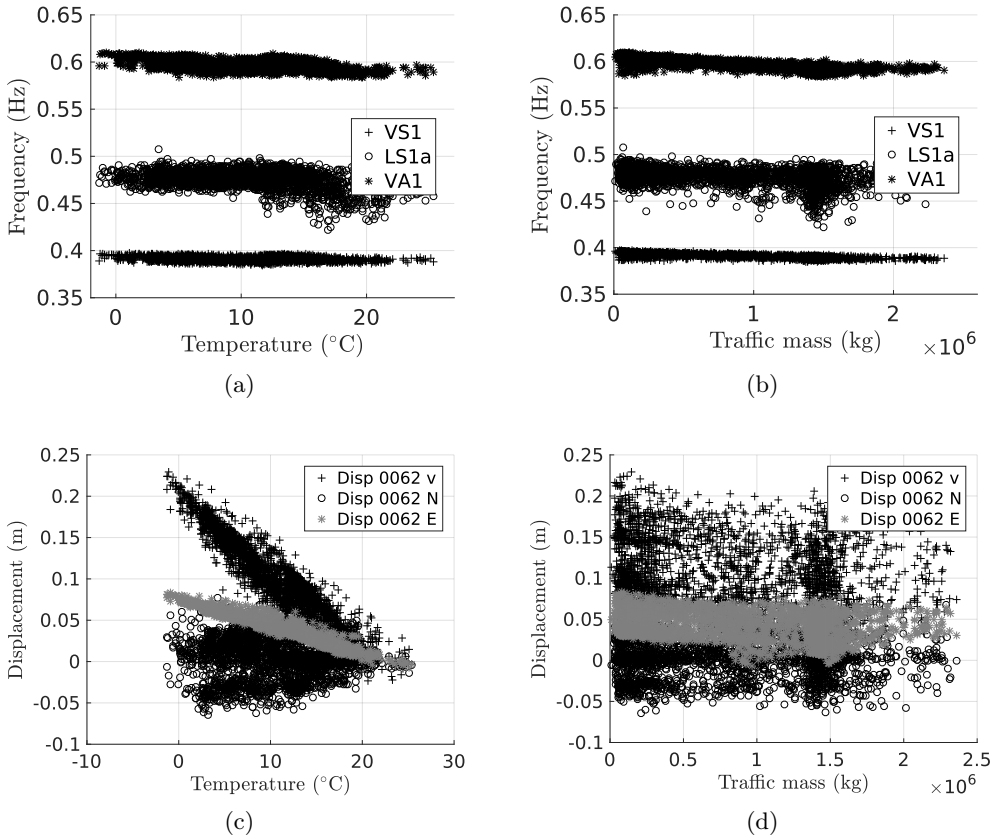


Figure 5.13: Post-processed data – May of 2009 to March of 2010 time period - natural frequencies (a), (b) and mid-span relative displacements (c) and (d).

5.3.2 Modelling of thermal and traffic effects

In this section, Tamar bridge FE model will be briefly described, along with to-be-identified structural parameters, and modelling aspects of its dynamic behaviour in the presence of traffic and thermal variations.

The FE bridge model has been developed using ANSYS Parametric Design Language (APDL) source code [203] and consists of approximately 45 000 elements, from which: expansion joints have been modelled with linear spring elements; truss members with fixed-rotation beams; deck/towers with shells and the cables and hangers with uniaxial tension only beam elements. The original FE model from Westgate has been altered in order to: simulate traffic in only one side of the bridge's

deck; and a modal assurance criterion (MAC) routine, which allows to obtain the correct mode shape for varying temperature/traffic, also had to be developed.

At first, it is important to discriminate the three parameters which will be identified. One is the stiffness of linear springs K_d, as seen in Fig. 5.14(b) and 5.14(d), which represents friction in the thermal expansion bearings of Saltash tower. The remaining two parameters are the initial strain ε_i in the two main and 16 stay cables of the bridge. For each cable type (main or stay), the initial strain is assumed constant along the cable length and across all cables. It is known that the simulated natural frequencies are sensitive to the cables initial strain parameters, as noted in Westgate and Brownjohn [204] analysis. In turn, the mid-span displacements are sensitive to the stiffness of the thermal expansion gap, as shown in Westgate et al. [205].

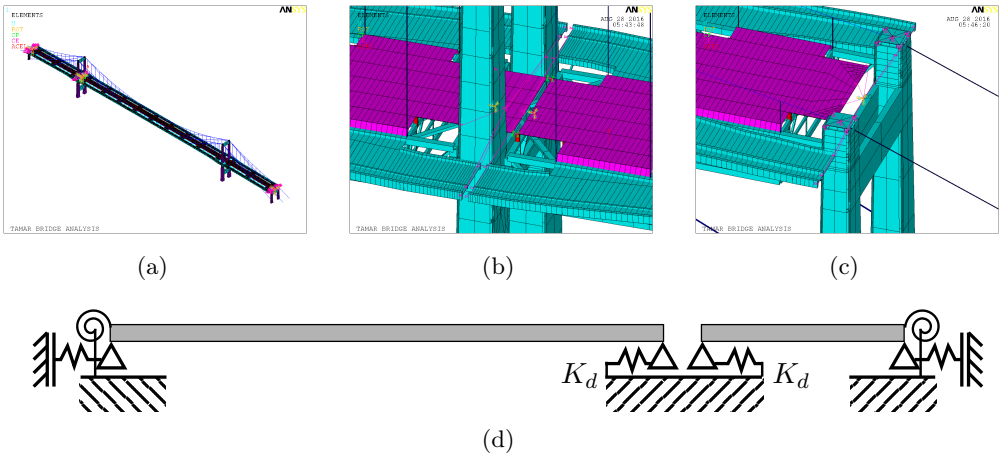


Figure 5.14: Tamar bridge FE model and detail of imposed constraints simulated as linear spring elements. Perspective view (a), expansion gap at Saltash tower (b), Plymouth embankment (c) and bridge boundary conditions diagram (d).

For future reference, it is worth mentioning some relevant information related with the bridge main and stay cables. Namely, the main suspension cables are made from 31 locked coil wire ropes, each 60 mm in diameter, and the overall diameter of the main cable is 380 mm, resulting in a cable cross-sectional area of 882.36 cm^2. The stay cables indicated in Fig. 5.12 have areas of 87.01 cm^2 for S2 and P2 (110 mm diameter strands) and 70.74 cm^2 for the remaning cables (102 mm diameter strands). Finally, Young's modulus of the cables was assumed as 155 GPa. The initial strain ε_i affects the axial tensile force N of the cables according to trivial constitutive laws, i.e. $N = E_c A_c \varepsilon_i$, where E_c, A_c represent the cables Young's modulus and cross-sectional area, respectively.

Secondly, and before detailing the modelling of thermal and traffic effects, the mode shapes and natural frequencies of the FE model (without applying any external effect) are shown in Fig. 5.15. It should be noted that certain mode shapes, such as the LS1b in Fig. 5.15(d), tended to disappear easily from the frequency spectrum when external effects were considered.

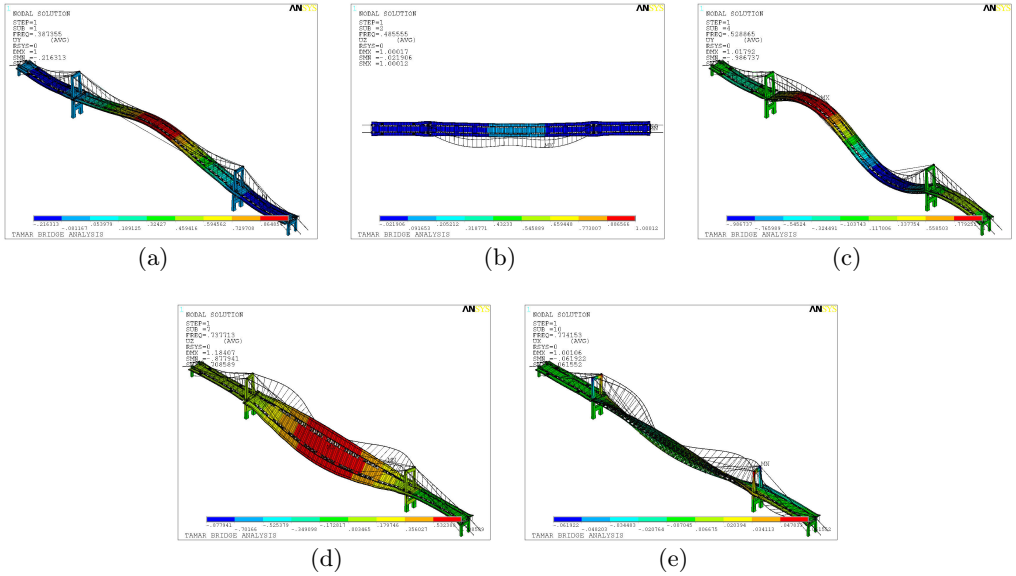


Figure 5.15: Tamar bridge FE model mode shapes, VS1 - 0.387 Hz (a), LS1a - 0.485 Hz (b), VA1 - 0.529 Hz (c), LS1b - 0.738 Hz (d) and TS1 - 0.774 Hz (e). See Section 5.3.1 for a reference of the frequency labels.

Thirdly, after having detailed the model parameters it is now necessary to highlight how temperature effects have been considered. For the present work, a regression was used to establish a simplified relation between temperature trends of a truss, deck, and cable temperatures of the bridge (sensor location can be seen in Fig. 5.12). These effects are considered as a static steady-state thermal analysis, with three different temperature sets: shaded elements, which represent the truss structure under the deck; the elements that represent suspension cables; and the lighted elements excluding the cables. The relation between the temperatures of the shaded, lighted and cable groups is assumed as

$$\tau_S = \begin{cases} 0.433\tau_c + 7.877 & \tau_c > 15 \\ \tau_c & \tau_c \leq 15 \end{cases} \qquad \tau_L = \begin{cases} 1.544\tau_c - 8.798 & \tau_c > 15 \\ \tau_c & \tau_c \leq 15 \end{cases} \tag{5.2}$$

where τ_c, τ_S and τ_L represent the temperature in the cable, shaded, and lighted elements, respectively. These linear relations are displayed in Fig. 5.16. It must be

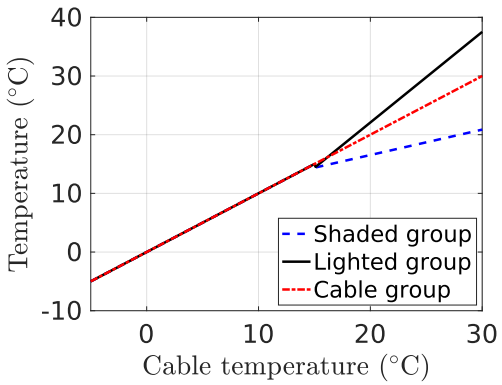


Figure 5.16: Assumed temperature relations between cable, shaded and lighted groups for Tamar bridge FE model.

stressed that the linear relations of Eq. (5.2) will only be applied when $T_c > 15°C$, which is thought as a notable change of cable temperature, where the temperature trends of the bridge components fork, otherwise to be applied as a uniform temperature T_c across all elements.

The aforementioned simplified thermal analysis has been detailed by Westgate in Section 4.4.3 [116]. Alternatively, a transient thermal analysis of Tamar bridge has also been developed by the same author, but presently only the simplified thermal analysis will be used to assess the FE model performance.

Fourthly, the effects of traffic load are also to be detailed. The traffic is assumed as a set of distributed mass points, evenly spread longitudinally across the bridge deck, and asymmetrically in the lateral direction, as shown in Fig. 5.17. The only remaining question is "how to compare this model against the monitored data?" According to Westgate et al. [205] in the Section "Traffic mass", the average number of vehicles *on the bridge* in the tolled direction can be assumed as approximately 1/43rd of the available half-hourly count. Hence, traffic effects were monitored as a function of a half-hourly total mass m_t, which was retrieved from the available toll-gate count, and in the FE model, the total mass is being fractioned into an half-hourly *on bridge* mass m_s, as $m_s = m_t/43$.

Finally, and gathering all the above mentioned information, the simulations of the FE model have been performed in a Latin hypercube space [200], where values of temperature, traffic mass and structural parameters were uniformly generated and the corresponding natural frequencies/displacements stored. A flowchart of the whole process is shown in Fig. 5.18. Initially, the natural frequencies were simulated and stored without any loading applied and with default values of structural

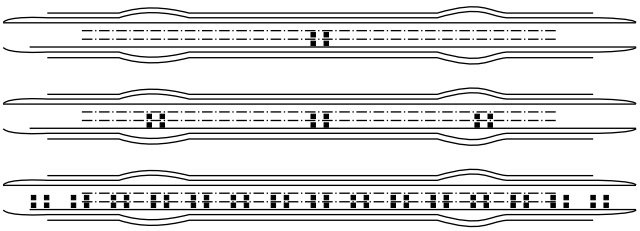


Figure 5.17: Traffic model of Tamar bridge FE model for 2 top, 6 middle and 30 vehicles, bottom. Each mass point represents half of the mass of a standard vehicle (1660 kg).

parameters. In the second phase, for each combination of loading $\boldsymbol{X}^m$ parameters $\boldsymbol{\Theta}$ the frequencies of each run were classified and stored, using the modal assurance criterion (MAC) with a fit of at least 80 %.

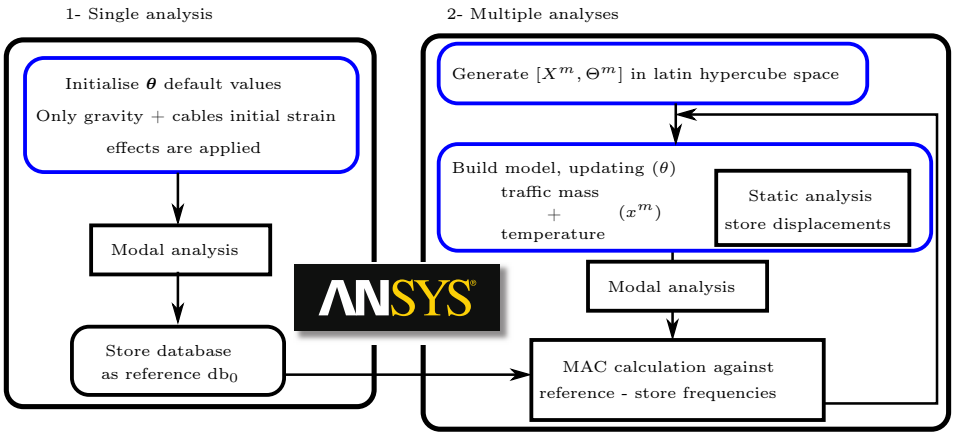


Figure 5.18: Simulation flowchart of the Tamar bridge FE model. Thermal and traffic effects and the calibration parameters have been updated in a Latin hypercube space.

5.4 Structural identification and discrepancy surfaces prediction

5.4.1 Modular Bayesian approach input dataset

Table 5.5 presents a summary of all the MBA input data, and the structural parameters $\boldsymbol{\theta} = \{\varepsilon_{iMC}, \varepsilon_{iSC}, K_d\}$ that have to be identified. Prior information of these parameters is considered as a uniform PDF, bounded by the intervals where the dataset $[\boldsymbol{X}^m, \boldsymbol{\Theta}]$ has been generated.

The mean and correlation functions of the mrGps are set as a polynomial regression $\boldsymbol{H}(\cdot) = 1$ and a linear correlation, respectively. The required computational

	Description
$\boldsymbol{\theta}$	initial strain of main ε_{iMC} and side cables ε_{iSC} and stiffness of linear springs at thermal expansion gap K_d
$\boldsymbol{X}^e$	bridge cable temperature and total mass due to traffic load from Plymouth-to-Saltash direction
$\boldsymbol{Y}^e$	natural frequencies determined by SSI and mid-span displacement from TPS
$[\boldsymbol{X}^m, \boldsymbol{\Theta}]$	combination set of $[\tau_c/m_t/\varepsilon_{iMC}/\varepsilon_{iSC}/K_d]$ generated in a latin hypercube space with τ_c [-5, 30]°C, m_t [0, 2.5×10^6]kg, ε_{iMC} [36.5, 2700]$\mu\varepsilon$, ε_{iSC} [36.5, 3700]$\mu\varepsilon$, and K_d [0, 10]kN/mm
$\boldsymbol{Y}^m$	simulations of natural frequencies/displacements for the $[\boldsymbol{X}^m, \boldsymbol{\Theta}]$ input

Table 5.5: MBA input dataset for Tamar bridge.

time was 6.38 h on an Intel Core i7-4790, 3.60GHz, 33 GB of RAM.

In order to sample the likelihood PDF with the Metropolis–Hastings algorithm, four Markov chains have been generated, with a standard multivariate normal distribution set as the proposal distribution. Total number of samples was 100 000 with a burn-in period of 3000 samples. Trace plots of accepted samples are shown in Fig. 5.19. These plots suggest that the chains are not correlated, and their convergence after the burn-in period has already become stable. The samples acceptance ratio was 39%.

In order to fully describe the results of the MBA the mrGp of the FE model and discrepancy function are discriminated in Tables 5.6, 5.7, 5.8, 5.9, 5.10, 5.11. As can be observed, eight responses have been approximated by the mrGps which is four times more responses than what Arendt et al.[146] have been able to approximate with their implementation (two responses). The core reason is the stability of the implementation conferred by the factorisation described in Section 3.7, which is an original contribution of this thesis.

Overall, the process variance of the vertical mid-span displacement of the bridge presented the largest variance for the FE model, discrepancy function and observation error, as can be seen in Tables 5.8, 5.10 and 5.11, respectively. These results correlate well with engineering knowledge, where it is known that the mid-span displacement of bridges is a highly informative measurement.

	τ_c	m_t	ε_{iMC}	ε_{iSC}	K_d
$\hat{\boldsymbol{\omega}}$	0.1	0.1	0.136	0.1	0.1

Table 5.6: Roughness parameters of the Tamar bridge FE model mrGp.

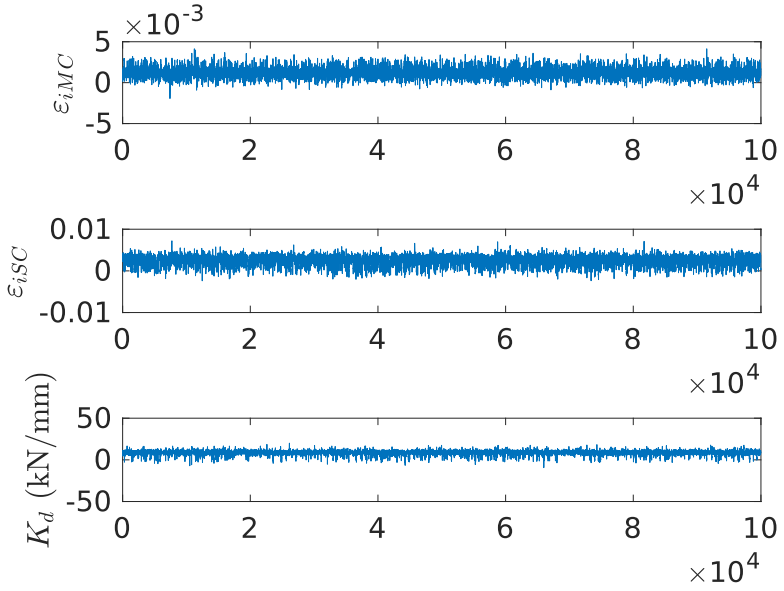


Figure 5.19: Samples trace plot of the identified structural parameters posterior PDF. Horizontal axes represent the sample number, and the vertical axes represent the structural parameters values in their respective units.

	VS1	LS1a	VA1	LS1b	TS1	0062 v	0062 N	0062 E
	Hz/(° C kg)					m/(°C kg)		
$\hat{\boldsymbol{\beta}}$	-0.068	-0.229	-0.071	-0.148	-0.012	-0.11	-0.023	0.035

Table 5.7: Regression coefficients of the Tamar bridge FE model mrGp.

(10^{-8})	VS1	LS1a	VA1	LS1b	TS1	0062 v	0062 N	0062 E
	(Hz^2)	(Hz^2)	(Hz^2)	(Hz^2)	(Hz^2)	(m^2)	(m^2)	(m^2)
$\hat{\boldsymbol{\Sigma}}^2$	211.0	16878	348.6	2165.1	1028.9	76.1×10^5	5.9	12569.0

Table 5.8: Process variance of the Tamar bridge FE model mrGp.

Furthermore, the regression coefficients of the discrepancy function seen in Table 5.9 are smaller for regressed displacements in comparison with the regressed natural frequencies of the bridge. Although additional verification is necessary, such fact indicates that the FE model predicts displacements more reliably than natural frequencies. Additionally, the roughness parameters in Eq. (5.3) attain large values, which indicates that there is a considerable amount of variability beyond what is induced by traffic and temperature. On the basis of the introductory text of

Section 5.3.1, it is reasonable to assume that such variability occurs due to wind.

$$\hat{\boldsymbol{\omega}}_{\tau_c,m_t} = \begin{pmatrix} 9.948 & 9.881 \end{pmatrix} \tag{5.3}$$

	VS1	LS1a	VA1	LS1b	TS1	0062 v	0062 N	0062 E
	Hz/(° C kg)					m/(°C kg)		
$\hat{\boldsymbol{\beta}}$	-0.295	-0.199	29.67	-0.532	-1.419	-0.504	6.041	-0.146

Table 5.9: Regression coefficients of the Tamar bridge discrepancy function mrGp.

(10^{-6})	VS1	LS1a	VA1	LS1b	TS1	0062 v	0062 N	0062 E
	(Hz^2)	(Hz^2)	(Hz^2)	(Hz^2)	(Hz^2)	(m^2)	(m^2)	(m^2)
$\hat{\boldsymbol{\Sigma}}^2$	4.43	1000.00	79.80	208.96	38.91	1471500	214.51	2200.00

Table 5.10: Process variance of the Tamar bridge discrepancy function mrGp.

(10^{-7})	VS1	LS1a	VA1	LS1b	TS1	0062 v	0062 N	0062 E
	(Hz^2)	(Hz^2)	(Hz^2)	(Hz^2)	(Hz^2)	(m^2)	(m^2)	(m^2)
$\hat{\boldsymbol{\Lambda}}$	4.7	1213.1	304.6	2672.4	1426.8	59.9×10^5	746.3	6025.3

Table 5.11: Observation error of the Tamar bridge discrepancy function mrGp.

5.4.2 Prediction of model discrepancy

Beforehand, it is important to predict a discrepancy function, whose information can be used to update the model, or added to the model output to compensate for inevitable modelling inadequacies. In the current section, some of the obtained discrepancy function predictions will be presented and commented. Since the outputs depend of temperature/traffic and mrGps are used for visualisation, the results assume the form of 3D statistical response surfaces, i.e. a mean 3D surface and a prediction interval cloud. However, for the sake of clarity the prediction intervals surrounding the mean surface have been omitted. Finally, it is important to mention that when the predictions of the model agree more closely to the monitored data, e.g. because the model has been fine-tuned and updated continuously, the closer the discrepancy function will be to a zero-mean uncorrelated Gaussian.

A first example of the mean discrepancy function of the natural frequencies of the Tamar bridge is shown in Fig. 5.20. The first thing to observe is that none of the mean surfaces is close to zero, and all display a correlated behaviour. As noted before, wind also affects the natural frequencies of the Tamar bridge, so it is plausible to assume that the visible correlation is due to wind. Secondly, some

patterns are visible in the lateral sway modes in Figs. 5.20(a) and 5.20(d), where the discrepancy increases smoothly with traffic mass up to a localised peak at 1400 ton. Note that the same peak occurs, irrespective of temperature, which indicates that it depends only of the modelled traffic effects and predominantly during rushing hours. Thus, it is reasonable to assume that it occurs because the developed model does not consider traffic from the Plymouth to Saltash direction.

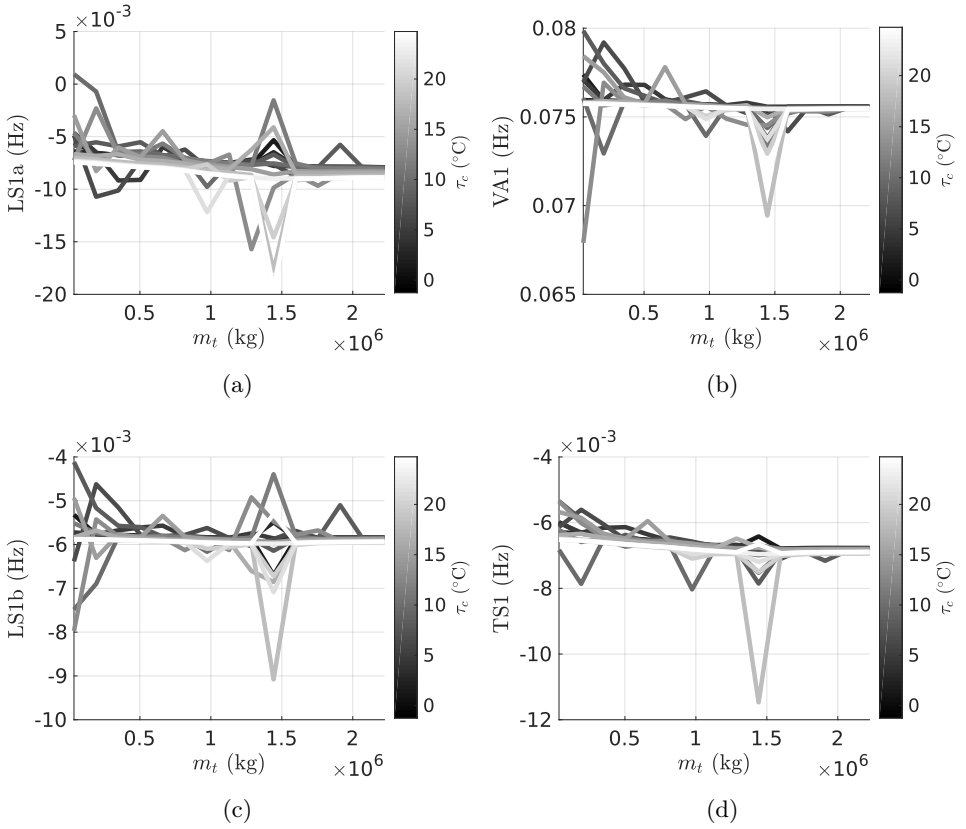


Figure 5.20: Prediction of discrepancy surface of the Tamar bridge natural frequencies LS1a (a), VA1 (b), LS1b (c) and TS1 (d) for varying temperature and traffic conditions.

Summarily, the results highlight the limitations of not considering a two-way traffic model, and assuming an asymmetric distribution of traffic mass. Equipped with this information, an analyst could integrate it in its model predictions, or carry out further updates.

Other example is shown in Fig. 5.21, where the model discrepancy of the mid-span displacements is being predicted by the MBA. The results indicate a better performance, since the shown surface resembles a zero mean uncorrelated Gaussian, particularly visible for the displacement in the lateral North direction. For example,

contrast Fig. 5.21(b) with the plots in Fig. 5.20. The former is much less smooth and regular, and its mean is very close to zero.

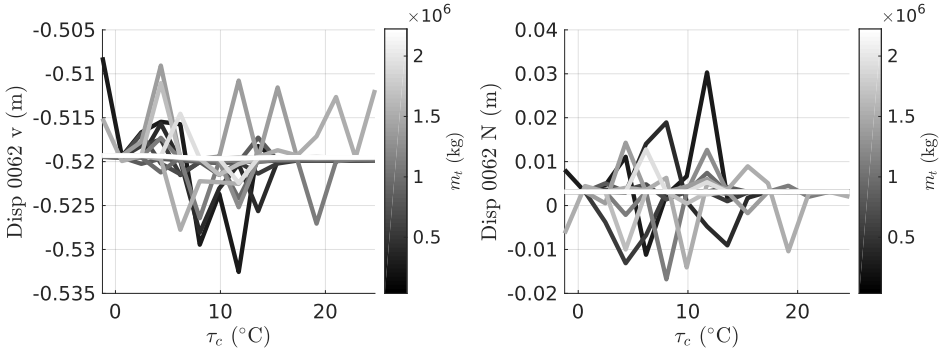


Figure 5.21: Predictions of discrepancy surface of the Tamar bridge mid-span displacements vertical (a) and North (b) for varying temperature and traffic conditions.

A final example is shown in Fig. 5.22. The shown plots are equivalent to Figs. 5.20(c) and 5.20(d) but with temperature in the abscissa. The model discrepancy for the two modes, LS1b and TS1, exhibits a temperature/traffic interaction, since recurrent peaks occur at different temperature and traffic values. It is important to always identify such patterns, and in which responses they become more proeminent, in order to ensure that the model predictions can be properly interpreted.

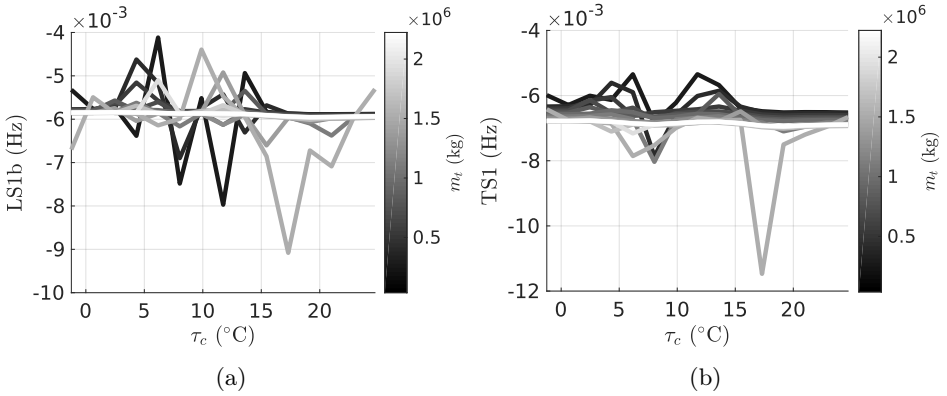
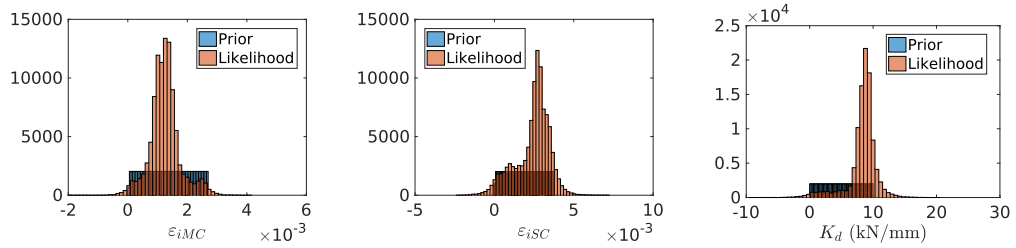


Figure 5.22: Predictions of discrepancy surface of the Tamar bridge natural frequencies LS1b (a) and TS1 (b) for varying temperature and traffic conditions.

5.4.3 Validation of identified calibration parameters

In this section, identified calibration parameters are presented and validated. Whereas the results of the previous section are useful for model updating, identifying the true value of calibration parameters is essential for damage detection or reliability analyses. It should be noted that compared to the hierarchical Bayes from Behmanesh et al., the MBA is unable to capture the inherent variability of calibration parameters. Therefore, note that the variance shown in the following results is associated with the attempt to estimate the parameters.

Sample histograms of the prior, likelihood are displayed in Fig. 5.23, and moments of the posterior distribution are shown in Eq. (5.4). Since an uninformative prior has been assumed, the posterior distribution and the maximum a posteriori (MAP) are proportional/equivalent to the values obtained from the likelihood.



(a) (b) (c)

Figure 5.23: Inference of calibration parameters: main (a) and sway cables (b) initial strain and stiffness of thermal expansion gap (c).

$$\mathrm{E}[\boldsymbol{\theta}|\boldsymbol{D}] = \begin{bmatrix} 0.0012 \\ 0.0024 \\ 8.3290 \end{bmatrix} \quad \boldsymbol{V}[\boldsymbol{\theta}|\boldsymbol{D}] = \begin{bmatrix} 0.25 & -0.081 & -168.47 \\ -0.081 & 1.17 & 519.28 \\ -168.47 & 519.28 & 5394863.84 \end{bmatrix} \times 10^{-6} \tag{5.4}$$

Subsequently, an attempt to validate these results is carried out. Opportunely, several expeditions which monitored Tamar's bridge behaviour at a local level, and other estimates from the calibration parameters, are available from past literature. Since the main cables have not been monitored in any way, and its strains vary along its length, more emphasis will be given to the stay cables and the thermal expansion gap.

The above estimates shall now be compared against a st-id confidence interval reported by Laory et al. [120] and Goulet [176]. An estimate of the cables internal forces was obtained based on the initial strain MAP and the FE model predictions for each type of cable (2 or 1,3,4). The information about each type of cable has

been detailed in Section 5.3.2. Results are displayed in Table 5.12. For the stay cable (SC) forces, there is a reasonable agreement between the two methodologies, since the MBA MAP falls near the upper limits of the model falsification confidence interval. However, for the stiffness of the thermal expansion gap, the MBA estimate is considerably less than the model falsification interval. A high value of this parameter would indicate that the bridge deck is prone to develop cracks. In order to investigate which estimate is closer to the true friction value, and to further analyse the stay cables behaviour, a comparison against monitored data will be detailed next.

Year	SC (2)	SC(1,3,4)	K_d	Method
	kN	kN	kN/mm	
2006	[674, 4045]	[548, 3289]	$[1\times10^4, 1\times10^{11}]$	model falsification
2009	3236	2631	8.32	modular Bayesian approach

Table 5.12: Identification of cable forces, stiffness of thermal expansion gap and comparison against model falsification.

The predicted stay cable forces are plotted along with monitored forces from existent strain gauge load cells, as shown in Fig. 5.24. It is also important to stress that the monitored data has not been used in any way to obtain the st-id posterior. Both histograms have been normalised in the coordinate axis by an estimate of probability density.

Firstly, it can be seen that the P3S cable has three to four times larger forces, and as reported in [211], has a stronger dependency on temperature, than other stay cables. It is not certain why such behaviour occurs. Secondly, it can be observed that the posterior obtained by the MBA falls between the P3S and the other cables histograms. Two reasons might aid to clarify this result:

1. Since the parameter is assumed constant across all cables, its posterior distribution falls between the two other histograms, acting as an average value.

2. the offset might occur because the parameter represents an initial strain, i.e. the strain relative to when the bridge cables have been installed on the bridge, and not relative to the strain existent when the strain gauge load cells have been installed.

Note that point number 1 assumes that the P3S cables force is genuine, in which case the posterior distribution would have to be shifted towards lower values (since there are much more cables in the lower cable forces region). Thus, it is more plausible to believe that point 2 is the underlying reason for the posterior position, and that the values it indicates are closer to the true structural behaviour of the cables.

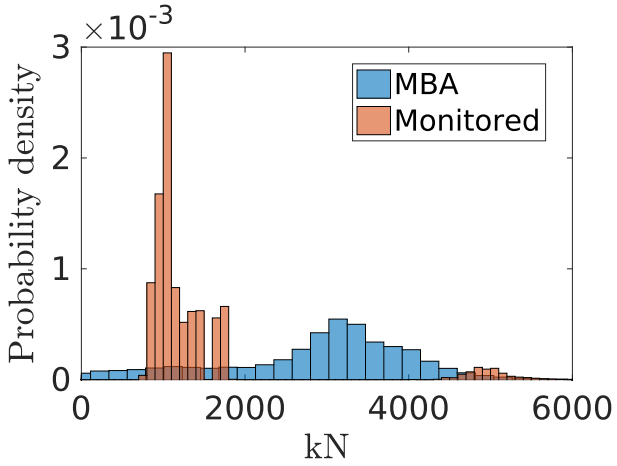


Figure 5.24: Tamar bridge histogram of monitored and identified stay cable forces (during January-July 2008). The right isolated peak represents the monitored P3S cable.

Finally, Battista et al. [212] recorded temperature and extension data in the thermal expansion gap, starting two months after the timeframe of data used for the MBA identification, i.e. in July 2010. Results from Battista's work revealed that the gap extension against temperature is perfectly adjusted to a linear relation (see Fig. 5.25), and does not indicate any relevant frictional force. If there was any friction, the empirical model in Fig. 5.25(b) would highlight some form of nonlinearity. Hence, a lower value of stiffness, such as the one identified by the MBA, suggests a better agreement with in-situ tests.

5.5 Summary

In this chapter MBA presented in Chapter 4 has been validated with two illustrative case studies for the purpose of st-id.

1. Measurements and simulations of strain and temperature of an aluminium truss bridge subjected to thermal expansion, have been analysed to identify the stiffness of its supports. Results have shown that this temperature st-id is feasible, and the MBA is able to accurately identify the true physical value of the support stiffness and the modal discrepancy of the FE model.

2. The MBA has been applied for structural identification of the Tamar long suspension bridge. The suspension cables initial strain and the stiffness of its thermal expansion gap were considered as model parameters to calibrate. A

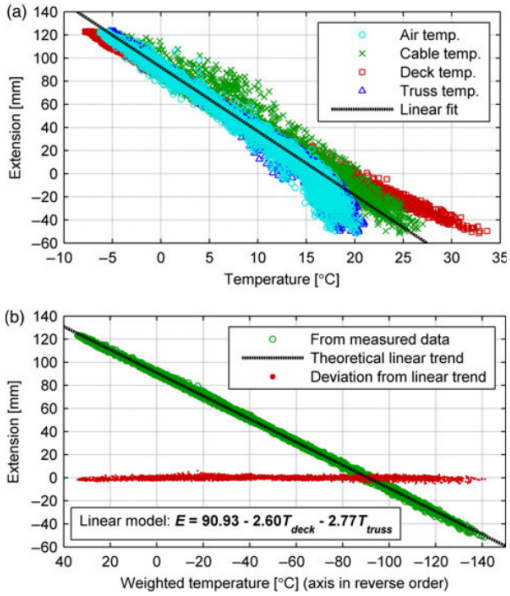


Figure 5.25: The deck extension plotted against: (a) the individual temperatures and (b) the combined deck and truss temperatures weighted according to a linear model, showing the improvement in accuracy achieved by using a derived empirical model. Figure reproduced from [212].

sophisticated FE model of the bridge, which considered traffic and temperature effects simultaneously was used for this purpose, and its performance has been assessed by the MBA through prediction of its model discrepancy.

Both of these examples have uncertainty due to observation error, model discrepancy and inherent variability of its parameters, which make the st-id challenging. Results indicate that:

- The MBA is able to effectively highlight trends and patterns of model discrepancy, as seen in Sections 5.2.4 and 5.4.2, which are essential for stake-holders and decision-makers who have to make a judgement based on a model's predictive ability. These results are unique relatively to previous SHM literature, which often assumed model discrepancy as an uncorrelated zero-mean Gaussian. This was the most significant challenge which had to be overcome when applying the MBA to the Tamar bridge.

- By the same token, the MBA is able to quantify how reliably environmental and operational variations effects are modelled, as seen in Sections 5.2.4 and 5.4.2, which has been acknowledged to be a major challenge in previous literature.

- Furthermore, the present results comprise the very first applications of the enhanced MBA to practical engineering applications.
- It is very important to consider the appropriate responses for the identification to be accurate and precise. This has been shown by the combinatorial analysis in Section 5.2.3, and will be detailed more clearly in Chapter 6.
- For the aluminium bridge example, the MBA was able to identify the true value of its supports stiffness with an accuracy five times superior to the one offered by a deterministic approach or other state of the art methodologies (such as the hierarchical Bayes approach). This has been shown by comparison of the relative error of the identified model parameter.
- In the Tamar bridge, the MBA identified the initial strain of its stay cables, and the offset between the estimate and monitored data was associated with the strain existent upon installation of the strain gauge load cells, and hence the estimated parameter is closer to the true structural behaviour of the cables.
- Similarly, the MBA estimated the stiffness of the bridge bearings, which according with monitored data should have minimal to no friction. The value estimated by the MBA is at least one-thousandth times smaller than the one estimated by Goulet's model falsification method.
- The FE model from Tamar bridge has been further developed, in order to simulate thermal and traffic effects simultaneously.
- Since Tamar bridge simplified steady-state thermal model provided sufficient insight into the structural parameters true value, a transient thermal analysis is not strictly necessary. Provided with this information, a structural analyst could save a considerable amount of time and effort devoted to modelling a certain phenomenon.

Chapter 6

Measurement system design

6.1 Introduction

This chapter describes a measurement system design methodology (MSD), developed to enhance the identifiability of the MBA. In contrast to Chapter 5, where st-id based on the MBA was performed from measured and simulated responses, MSD requires only simulated responses to achieve the best measurement system configuration.

Generally speaking, the ultimate goal of MSD is to establish the optimal sensor configuration of a monitoring system, in order for it to comply with the interests of decision-makers and stake-holders. Imposed constraints will range from economic limitations/storage capabilities, accessibility of certain locations of the infrastructure and energy consumption of the sensory system, amongst others. On the other hand, their interests will usually be focused on performance criteria which are not-directly measurable, e.g. reliability of the monitoring system, the reserve capacity that a structure maintains, or minimisation of Type I and II errors for damage detection. Such performance criteria are usually obtained from analysis of identified physical properties: natural frequencies [213], initial strain in a bridge suspension cables, etc. Finally, MSD will probably remain as an open-ended problem, since it is not feasible to consider *every* possible sensor configuration. Instead results typically indicate the best sensor configuration out of a set of available key locations/ options. It should also be noted, that most of current SHM practice in this subject relies on specialist experience, rather than solid theory, often introducing biases and/ or redundant information, etc.

Similarly to st-id, probabilistic MSD methodologies which consider uncertainties more comprehensively and more rigorously are expected to have a relative upper edge. Several Bayesian methodologies have already been developed for this purpose,

with some form of information related with model parameters posterior PDF typically selected as a performance criterion. Examples include the norm (determinant or trace of the Fisher information matrix [214]), information entropy [215], Bayes risk, and more recently, the identifiability of model parameters [174].

Henceforth, the objectives of this chapter are three-fold:

- present a Fast Preposterior Analysis (FPA), able to estimate the best sensor configuration which maximises identifiability of the enhanced MBA, based on a preposterior covariance criterion (Section 6.2). Since the MBA developed in this thesis is unique, so it is its MSD counterpart;
- provide a broader view on the topic of identifiability of the MBA, in order to understand what are the responses which maximise its identifiability, and how can the developed methodology be used to estimate such responses. Note that there is not yet sufficient insight on these questions from previous research;
- and demonstrate its feasibility for a simulated cantilever beam example, already presented in Section 4.7, and the thermally expanded aluminium bridge case-study of Section 6.4. Validation is carried out by comparing the results against a st-id combinatorial analysis, when monitoring data is made available (Section 6.4.1).

Finally in Section 6.5, a short summary of the current chapter is presented.

6.2 Fast preposterior analysis

6.2.1 Original preposterior analysis

Having presented the formulation of the multiple response MBA in Chapter 4, it is now possible to investigate the question of which responses should be chosen in order to improve the identifiability of the method, i.e. to make sure that the identified $\boldsymbol{\theta}$ in Eq. (4.4) attain their true physical value, and that $\boldsymbol{\delta}$ highlights the true inadequacies of the model. Note also, that in contrast with the FPA developed by Arendt, the implemented methodology allows to estimate the best sensor configuration for st-id of multiple calibration parameters.

To this end, the preposterior analysis:

- establishes a prior knowledge on any unknown quantities of Eq. (4.4) (hyper-parameters of mrGp, structural parameters);
- generates realisations of these quantities in a Monte Carlo sampling space;

- and takes an average across results of the structural parameters posterior. The identified covariance matrix of the posterior probability density function (PDF) is hereafter called preposterior covariance, in the sense that it is based on prior knowledge. This is the performance criterion which has to be minimised in order to achieve better identifiability with the MBA.

A final note is that if there is any prior information, or any heuristics relatively to the most relevant responses, they should be incorporated in any of the variables in step 1. Thus, the core optimisation of the preposterior analysis can be stated as

$$[\boldsymbol{Y}, \boldsymbol{X}]_{\text{best}} = \underset{\boldsymbol{D}^m}{\text{argmin}}[\hat{\Sigma}^2_{\theta,pp}], \tag{6.1}$$

where $\boldsymbol{D}^m = \{\boldsymbol{X}^m, \boldsymbol{\theta}^m, \boldsymbol{Y}^m\}$ represents a simulation dataset and $\hat{\Sigma}^2_{\theta,pp}$ represents the preposterior covariance metric. In essence, the prior information and simulated datasets are supplied to the main calibration equation and the variance of the posterior PDF of the structural parameters is obtained from the main equation as follows

$$\boldsymbol{Y}^e(\boldsymbol{X}^e) = \boldsymbol{Y}^m(\boldsymbol{X}^e, \underbrace{\overbrace{\boldsymbol{\theta}^*}^{prior}}_{\hat{\Sigma}^2_{\theta,pp}}) + \overbrace{\boldsymbol{\delta}(\boldsymbol{X}^e)}^{prior} + \overbrace{\boldsymbol{\varepsilon}}^{prior} \tag{6.2}$$

Referring to Fig. 1 of Arendt publication or Fig. 6.1 (flowchart on the right), the process of estimating $\hat{\Sigma}^2_{\theta,pp}$ is broken into eight steps enclosed in three groups: preliminaries, Monte Carlo Loop and calculation of this metric. Of these three groups, the Monte Carlo Loop is the most computationally intensive one, since it requires over each iteration: the calculation of the posterior distribution of the computer model in step 6; the estimation of the discrepancy function hyperparameters also in step 6; and the calculation of the posterior covariance of $\boldsymbol{\theta}$ in step 7.

6.2.2 Reuse of estimated mrGp hyperparameters

This section details a numerical strategy to minimise the computations required during the Monte Carlo (MC) loop of the original preposterior algorithm. A flowchart of this approach is shown in Fig. 6.1.

It is divided into two phases, which are then subdivided into three steps.The first phase can be thought of as a preliminary run, with all sensors active, hence all available q responses are being considered. The discrepancy function $\boldsymbol{\delta}$ mrGp hyperparameters, $\boldsymbol{\phi}_j^{(i)}$ with $j = 1 \ldots q$, are saved during this phase, at each iteration i of the Monte Carlo (MC) loop. The second phase is a combinatorial or optimisation run, similar to what has been presented by Arendt et al., with the particularity

that the required hyperparameters of $\boldsymbol{\delta}$ do not have to be computed. Instead, the corresponding hyperparameters $\boldsymbol{\phi}_k^{(i)}$ for each sensor configuration are reused from what has been stored in the first phase. A comparison with Arendt approach will be presented in Section 6.4.1.

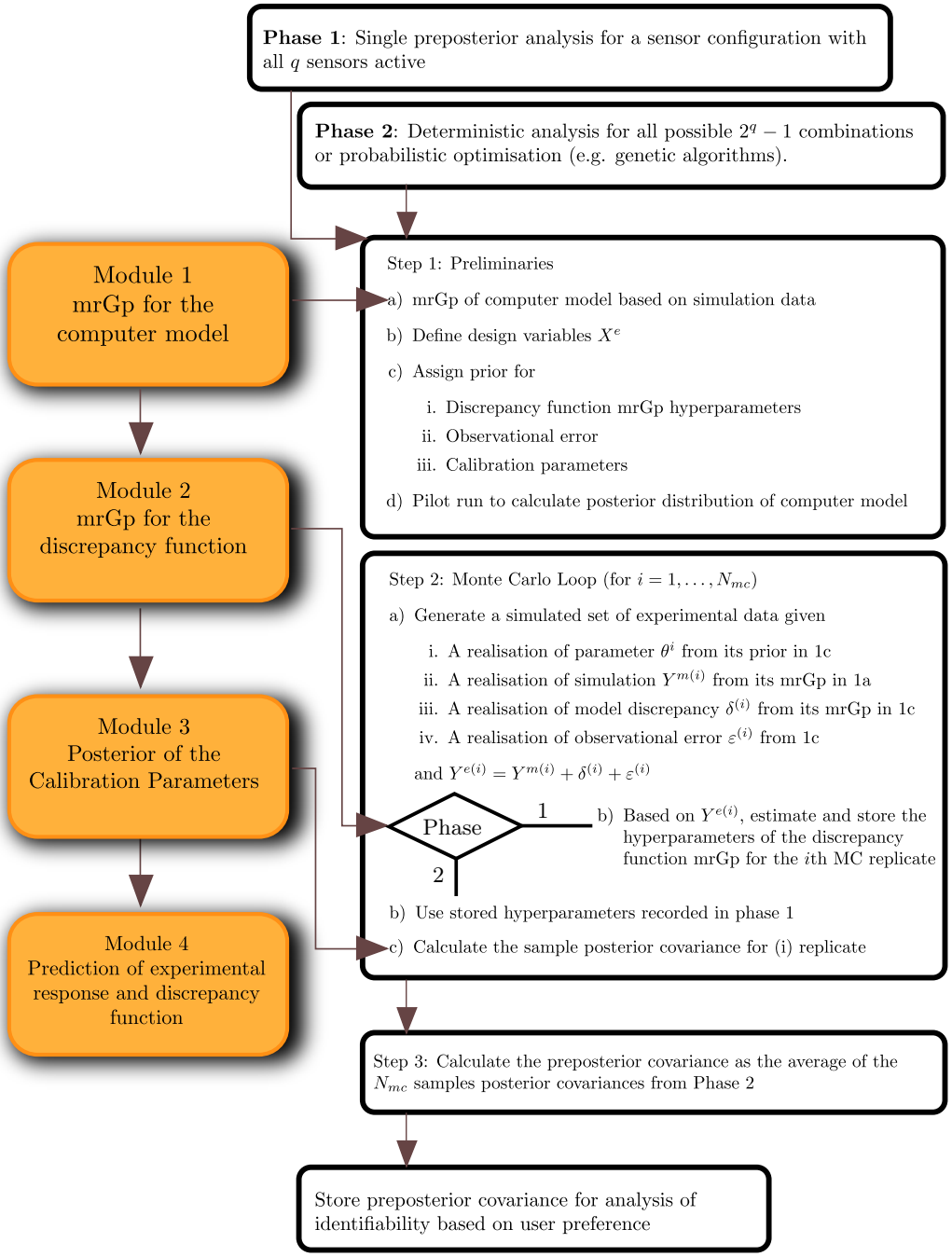


Figure 6.1: Flowcharts of: MBA on the left and FPA analysis on the right.

It is also necessary to choose how to search for the sensor configuration that yields the minimum $\hat{\Sigma}^2_{\theta,pp}$. This can be either a combinatorial analysis with $2^q - 1$ runs or an optimisation routine, and depends heavily on the number of sensors under consideration. As the number of sensors q grows, the computational cost of the combinatorial analysis becomes prohibitive, and optimisation algorithms, such as genetic algorithms, are better suited to estimate the optimal sensor configuration.

The FPA algorithm shall now be fully described:

Fast Preposterior Analysis algorithm

PHASE 1:

In this phase all q responses for all available sensors are considered in the analysis. Move to Step 1.

PHASE 2:

In this phase a combinatorial or optimisation routine takes place, depending of the objective at hand. If a combinatorial analysis is run, iteratively set different sensor configurations out of the $2^q - 1$ available combinations. Otherwise minimise the preposterior covariance as a function of the sensor configuration. Stop if any convergence criterion is met. Otherwise move to Step 1.

STEP 1

In this step the preliminary computations do take place:

a) A mrGp of the computer model is fitted on available simulation data, similarly to module 1 of the MBA. Hyperparameters of the mrGp $\boldsymbol{\phi}_m = \{\boldsymbol{\beta}_m, \boldsymbol{\Sigma}^2_m, \boldsymbol{\omega}_m\}$ are estimated.

b) Define the experimental design variables $\boldsymbol{X}^e$, which should be in the same range as the available simulations $\boldsymbol{X}^m$.

c) Assign prior PDFs individually for each of the following terms:

i. Discrepancy function mrGp hyperparameters $\boldsymbol{\phi}_\delta$.

ii. Observational error $\boldsymbol{\varepsilon}$.

iii. Structural parameters $\boldsymbol{\theta}$.

d) In this step, based on the mrGp fitted in step 1a) and the prior of the structural parameter $\boldsymbol{\theta}$ set in 1c)iii, the posterior distribution of the mrGp computer model at $\boldsymbol{X}^e$ is calculated.

STEP 2

In this step a Monte Carlo Loop for $i = 1, \ldots, N_{mc}$ is run, upon which the following sub-steps take place:

a) Generate a simulated set of experimental data $\boldsymbol{Y}^e$ given:

i. A realisation of parameter θ^i from its prior in sub-step 1c) iii.

ii. A realisation of simulation $Y^{m(i)}$ is generated from its mrGp in 1a) at $(\boldsymbol{X}^e, \theta^{(i)})$.

iii. A realisation of model discrepancy $\delta^{(i)}$ without any data is generated from its mrGp in sub-step 1c).

iv. A realisation of observational error $\varepsilon^{(i)}$ from 1c).

and finally, combine the above elements as $Y^{e(i)} = Y^{m(i)} + \delta^{(i)} + \varepsilon^{(i)}$.

b) In this sub-step similar to module 2 of the MBA, a Gaussian process approximates the discrepancy between experimental data and the posterior mrGP which was computed in sub-step 1d). If phase 1 is being run: based on $Y^{e(i)}$, the hyperparameters of the discrepancy function $\boldsymbol{\phi}_\delta = \Sigma_\delta^{2(i)}, \omega_\delta^{(i)}, \varepsilon_\delta^{(i)}$ mrGp for the ith MC replicate are estimated and stored for posterior use; if in phase 2 instead, use the hyperparameters $\boldsymbol{\phi}_\delta$ previously stored, corresponding to the current combination of sensors under analysis.

c) Similarly to module 3 of the MBA, and based on the hyperparameters $\boldsymbol{\phi}_m, \boldsymbol{\phi}_\delta$ and simulations/generated data of step 2a), calculate the sample posterior covariace of θ for replicate i.

STEP 3

In this step the calculation of the preposterior covariance matrix and its norm take place, based on the N_{mc} available samples as

$$\hat{\Sigma}^2_{\theta,pp} = \left\| \frac{1}{N_{mc}} \sum_{i=1}^{N_{mc}} \boldsymbol{V}^i[\boldsymbol{\theta}|\boldsymbol{Y}^{e(i)}, \boldsymbol{Y}^m] \right\|, \tag{6.3}$$

where $\boldsymbol{V}^i$ denotes the posterior covariance matrix of structural parameters $\boldsymbol{\theta}$ for iteration i. Return to Phase 2.

In the next sections, application of this algorithm to a case-study of a simulated cantilever beam and a reduced-scale aluminium bridge shall be presented, to illustrate its performance.

6.3 Simulated cantilever beam

In this example a cantilever beam subjected to a point end load suffers a deflection, which is being simulated at the tip, mid and a quarter of the beam span. Other simulated quantities include the rotation angle of the tip of the beam and its deformation energy. For notation purposes, these quantities will be denominated by a letter as presented in Table 6.1.The simulated input dataset of these responses has 54 data points in the Young's modulus and applied force range, as specified in

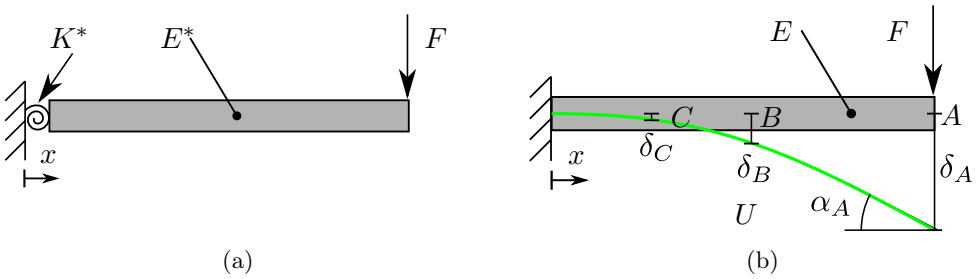


Figure 6.2: Cantilever beam example. True cantilever beam (a) and Idealised cantilever beam 6 responses (b).

Table 4.1. For validation purposes the experimental dataset has 25 data points.

	Observed response
A	Displacement at point A
B	Deformation energy of beam
C	Rotation angle at point A
D	Displacement at point B
E	Displacement at point C

Table 6.1: Notation of monitored responses for cantilever beam example.

The mean and correlation functions of the cantilever beam mrGps are set as a constant value of polynomial regression $\boldsymbol{H}(\cdot) = 1$ and a linear correlation function, respectively as described in Tables 3.1 and 3.2, whereas the structural parameter, the discrepancy function and observation error hyperparameters prior information are considered as a uniform PDF, in the intervals specified in Table 6.2. It should be noted that the variances Σ_δ^2 and λ are generated matrices, on which all entries will be scaled by the minimum and maximum factors presented in the table.

	$\boldsymbol{\theta}$ structural parameter	$\boldsymbol{\delta}$ mrGp hyperparameters			$\boldsymbol{\varepsilon}$
	E (MPa)	ω_δ	Σ_δ^2	β_δ	λ
min	20×10^3	0.1	1×10^{-2}	-5	0
max	100×10^3	9.0	2×10^1	5	1×10^{-4}

Table 6.2: Prior information of structural parameter, discrepancy function mrGp and observation error for cantilever beam example.

6.3.1 Results of cantilever beam optimal sensor configuration

In this section validation of the FPA with the cantilever beam simulated example is presented. The results of the optimal sensor configuration will be compared against a subsequent st-id performed with the MBA.

Using the prior information displayed in Table 6.2, Fig. 6.3 displays the resulting preposterior covariance $\hat{\Sigma}^2$ of the cantilever beam Young's modulus, which has been estimated with the FPA, against the relative bias $\epsilon = \frac{|K - \mathrm{E}[\theta]|}{K}$ that a st-id performed with the MBA attained. Ideally, the clusters of responses seen in this

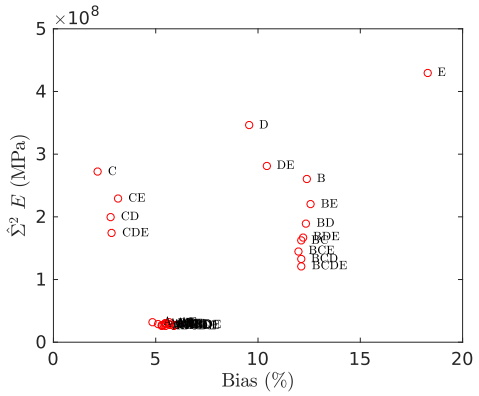


Figure 6.3: Relation of preposterior covariance vs bias for cantilever beam example.

plot should follow a linear trend. Failure to do so indicates the presence of Type I or Type II errors. In the current context these errors imply:

- Type I: false positive. A combination of responses which has a low preposterior covariance but has high bias.
- Type II: false negative. A combination of responses which has a high preposterior covariance but the bias is low.

Since the rotation of the beam tip, C cluster, is at the left top side of the plot, it indicates a Type II error, i.e. a failure to detect a good sensor configuration candidate for st-id. It should be noted that the displacements along the beam, clusters A, D and E do follow reasonably well a linear trend, whereas the deformation energy and the rotation do deviate, although the former not significantly.

In practice, an analyst which intends to design a measurement system will be uncertain regarding the whereabouts of the structural parameters and the extension of the model discrepancy. However, in Fig. 6.3 the preposterior covariance is able to classify response A as more informative than D and E, which is related to how much the response varies for the same variability of the Young's modulus. The simulated data and the uninformative prior knowledge, suffices to capture this trend.

In contrast, the end-rotation of the beam is classed as much less informative than what it actually is. In the author's opinion, this occurs because the rotation is the response which is most affected by model discrepancy. Since the discrepancy

is induced by the rotational spring, as seen in Fig. 6.2(a), the rotation is the effect that is more directly related to it, and the supplied uninformative prior knowledge is insufficient to capture this effect.

To further justify this reasoning, it is elucidative to examine results of application of the MBA and preposterior analysis by other authors, such as Jiang et al. [148] and Arendt et al.[146]. In their simply supported beam example, combinations including the beam's plastic dissipation energy brought a considerable improvement to the st-id, as seen in Tables 2 and 3, respectively for Jiang and Arendt. Since the model discrepancy in their example is the plastic material law of the beam, it seems reasonable that the plastic dissipation energy response would be significantly affected by it.

Subsequently, the prior information presented in Table 6.2 is to be improved on the basis of what has been described above. Basically, a sufficiently experienced modeller would have to guess which responses are to be more affected by modelling assumptions and simplifications, and would convey this information through the prior knowledge supplied to the preposterior analysis. To illustrate this concept, in the present example the prior information of the process variance of the model discrepancy Σ_δ^2 will be changed. This hyperparameter is a positive-definite matrix with dimensionality as the number of considered responses, i.e. $q = 5$ for the present case. The previously reported results had the same maximum and minimum variance for all response entries. Herein the maximum variance for the rotation angle at the beam-end response shall be reduced from 20 to 0.1.

The resulting bias preposterior covariance relation is presented in Fig. 6.4.As can be seen, the cluster of the beam rotation angle is now accurately classed as a response which provides good identifiability. However, combinations of this response with others are also affected and their preposterior covariance became lower, creating some Type I errors. Future developments of the FPA should try to minimise such spurious effects.

These results suggest that, prior information of the discrepancy function hyperparameters is essential in order to accurately classify informative responses for st-id using the MBA. For the considered example, this was achieved by improving the model discrepancy of a response dependent on the rotational spring. It is also proposed, that responses which have a high dependency of the structural parameters $\boldsymbol{\theta}$ and the model discrepancy $\boldsymbol{\delta}$ tend to be more informative. This is reasonable in light of the terms of the main calibration equation, Eq. (6.2).

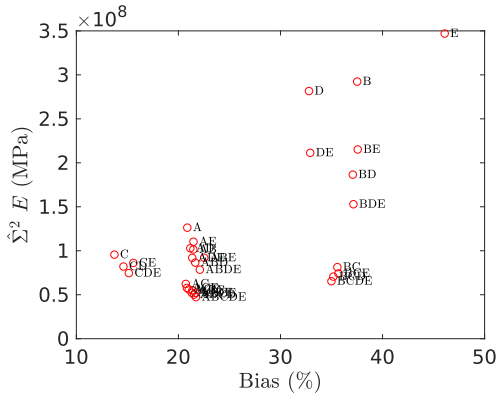


Figure 6.4: Relation of preposterior covariance vs bias for cantilever beam example with an informative prior.

6.4 Aluminium bridge subjected to thermal loading

The second case-study for validation of the FPA is the reduced-scale laboratory aluminium bridge, which has already been detailed in Section 5.2. A numerical model of the structure was also developed to study the phenomena. The stiffness of a pair of springs located at one of the ends of the bridge will be considered as a structural parameter to be identified. This inference by the MBA is not the primary goal of this work, but it shall be used as a validation tool to the optimal sensor configuration determined by the FPA. Nine readings/simulations of strain are used as input to the MBA and FPA methodologies, and shall be detailed in this section as well. Section 6.4.1 details results of application of the MBA and FPA.

6.4.1 Results and discussion of the FPA and its validation

In this section the results of the FPA and its validation against the st-id performed by the MBA are shown. The methodology was applied for all the nine available strain measurements sites seen in Fig. 5.4, based on a combinatorial analysis with all the available configurations, i.e. $2^q - 1 = 511$ runs.

Considerations of the analyses and computational effort

Having established the responses from the experiment and model presented in Section 6.4, non-informative uniform priors have been assigned to the quantities described in Step 1c) of the FPA, and are shown in Table 6.3. This decision is justified by an intent to test a worst case scenario, when there is no prior information available. For a definition of the mrGp hyperparameters see Sections 3.2 and 3.3. The

functions in the mean and covariance of the mrGps are set according to Tables 3.1 and 3.2 as a polynomial regression $\boldsymbol{H}(\bullet) = 1$ and a linear correlation function, respectively.

	$\boldsymbol{\theta}$ structural parameter	$\boldsymbol{\delta}$ mrGp hyperparameters			$\boldsymbol{\varepsilon}$
	K (N/mm)	ω_δ	Σ^2_δ	β_δ	λ
min	300	1×10^{-2}	1×10^{-8}	-2	0
max	900	1×10^{1}	1×10^{1}	2	1×10^{3}

Table 6.3: Prior information of structural parameter, discrepancy function mrGp and observation error for aluminium bridge example.

Finally, the mrGps require a set of input training data of temperature/strain $\boldsymbol{X}, \boldsymbol{Y}$. For the FPA, this data consisted of 14 simulations across all the nine responses, in the range visible in Fig. 5.5(a). For the MBA, 36 measurements were also added as an input to train the discrepancy function mrGp.

Given the above information, the Monte Carlo loop of the FPA routine was set to take 100 samples in the above-defined prior space. Justification for this amount of samples against a larger number (1600), as employed by Jiang et al. [148], is that the convergence of the preposterior covariance $\hat{\Sigma}^2$ became stable relatively quickly, as can be seen in Fig. 6.5.

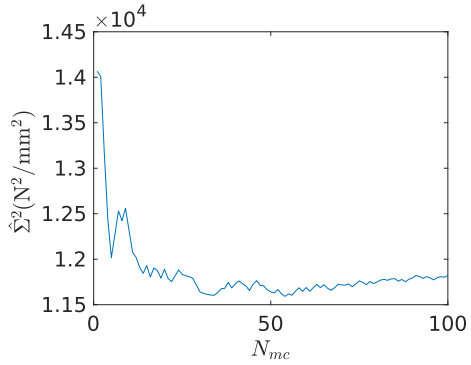


Figure 6.5: Convergence of estimated preposterior covariance during Monte Carlo loop, for a configuration with five sensors.

The computational effort of the whole combinatorial analysis was approximately six hours on a single-processor, 3.6 GHz, i7 machine. Arendt et al. [147] (Conclusions section) reported the computational effort of the original preposterior analysis on a simply supported beam simulated example, as 0.018 minutes for each MC sample, on a single-processor, 3 GHz, i7 machine.

Although these are different problems with a different data size a comparison might be appropriate, since both algorithms standardise input data, and in the above

publication data size was reported to not have a significant influence in the computational effort. Considering these assumptions to establish a comparison, the preposterior analysis proposed by Arendt et al. would take $0.018 \times 100 \times 511/60 = 15.3$ hours for this example, which is roughly three times what the FPA took. Unfortunately a similar comparison with Jiang et al. [148] surrogate preposterior method cannot be provided, since no information on the computational effort was reported.

Finally, the MBA was also run for all possible combinations, and the measurements from the experiment have been effectively used to identify the properties of the spring stiffness PDF (mean value $\mathrm{E}[\theta]$ and its posterior variance). The relative bias $\epsilon = \frac{|K - \mathrm{E}[\theta]|}{K}$ was also calculated. The same data size, prior structure for K and regression/correlation functions as the FPA were considered.

Optimal sensor configuration based on the FPA

This section presents the results of the application of the FPA and the MBA methodologies to the aluminium bridge case study. Fig. 6.6 displays the relation between the relative bias ϵ against the posterior variance, that were achieved by the MBA combinatorial analysis. As can be seen there is a clear tradeoff between these two quantities, which is a common scenario when using an estimator such as Bayesian inference, see for example Section 9.6 of Scharf [216]. This result counters the claim from Arendt et al. that, a decrease in the variability of the posterior PDF improves the identifiability of the solution. The question then is, whether it is possible to quantify this tradeoff with the FPA.

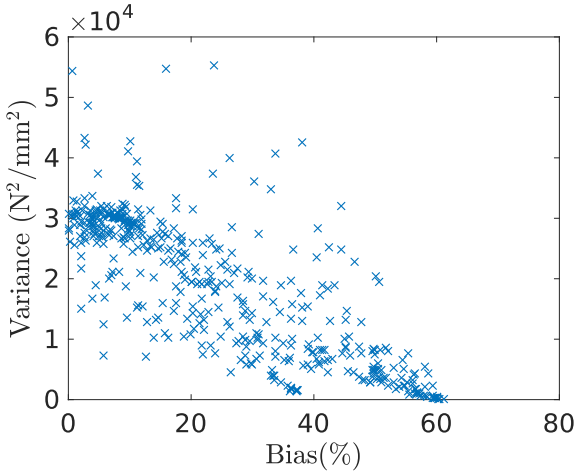
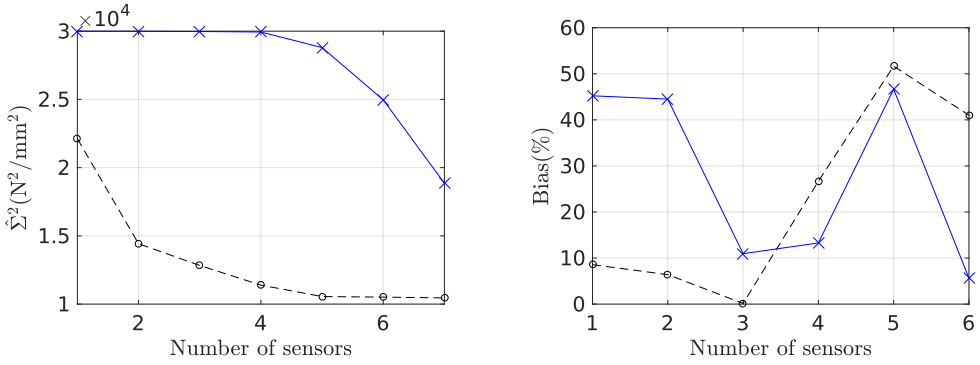


Figure 6.6: Bias-variance tradeoff observed for all sensor configurations (crosses) of the MBA combinatorial analysis.

Subsequently, the sensor configurations analysed by the FPA were grouped according to their number of sensors, and Fig. 6.7(a) shows the minimum and maximum values of preposterior covariance for each of these groups. In general both the maximum and minimum values tend to decrease with an increasing number of sensors. At this point it is useful to compare, how the percentual bias that these sensor configurations attain (calculated with the MBA) behaves at these maximum/minimum values of $\hat{\Sigma}^2$, which can be seen in Fig. 6.7(b). It is important to mention that these figures display a relationship between two outputs of the two algorithms for the same sensor configurations, and should be analysed collectively.

The difference between the maximum and minimum values of $\hat{\Sigma}^2$ is larger in the middle region (3, 4 and 5 sensors), and an inverse relation (shorter difference) exists relative to their equivalent bias in Fig. 6.7(b). More importantly, it is noticeable that in this central region, the bias increases as the number of sensors increase. Therefore both the minimum and maximum preposterior covariances should be examined carefully to locate this central region.



(a) (b)

Figure 6.7: a) Minimum/maximum values of the preposterior covariance obtained by the FPA (dashed line with circles and solid line with crosses, respectively) b) Percentual bias obtained from the MBA for the sensor configurations displayed in Fig. 6.7(a).

Notice also that both methodologies FPA/MBA were unable to recover a solution with a number of sensors above eight/seven, respectively, mainly due to numerical stability issues. This is related to the conditioning of the mrGp covariance matrices, and depends of factors such as the type of correlation function assumed, or how the sampling has been performed (grid-sample or as a uniform random Latin-hypercube). Similar issues were reported by Arendt et al. [146] in the Conclusions section.

Examining the sensor configurations with minimum preposterior covariance in this region, Fig. 6.8 and Table 6.4 highlight their placement on the actual structure and moments of the posterior PDF identified by the MBA, respectively. Note that similarly to what has been presented in Section 5.2.3, measurements of strain at locations B and G are selected as more informative. It is recalled that for this problem, one of the sources of model discrepancy is the difference in temperature between the bottom and top of the bridge. Thus, a sensor configuration which saturates at the centre of the bridge and between its top and bottom, is more sensitive to a change of the bridge supports stiffness or the discrepancy between the model and the actual bridge. Since the configurations shown in Fig. 6.8 follow these two rules, it is plausible that they are able to correctly identify the spring stiffness value. These configurations have measurements at the top and bottom of the bridge and

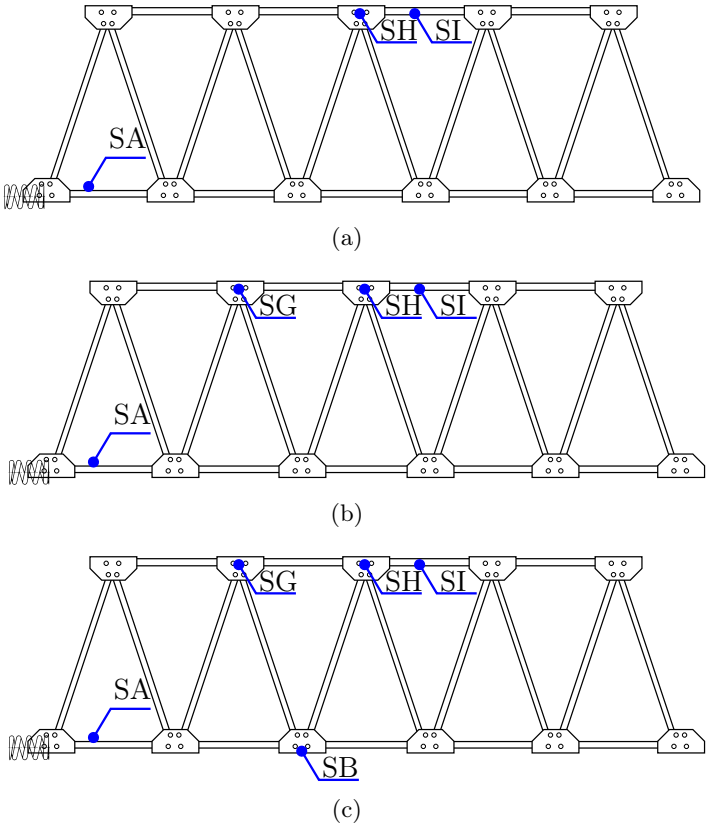


Figure 6.8: Diagrams of best sensor configurations obtained by the FPA for three (a) four (b) and five sensors (c). Strain gauges are classed as SA-SI.

almost exclusively at its centre (sensor A is the exception). As the number of sensors increases, the remaining bottom/top locations (sensor B and G) are also being selected.

	A	B	C	D	E	F	G	H	I	E[θ] (N/mm)	$\sigma[\theta]$ (N/mm)
3	x							x	x	551.47	168.49
4	x						x	x	x	699.55	155.96
5	x	x					x	x	x	837.52	59.44

Table 6.4: Best solution for 3, 4 and 5 sensors - based on the preposterior covariance minimum.

Figure 6.9 complements the information in Table 6.4, displaying the prior and posterior PDFs of the spring stiffness inference. Just as in Fig. 6.6 a tradeoff

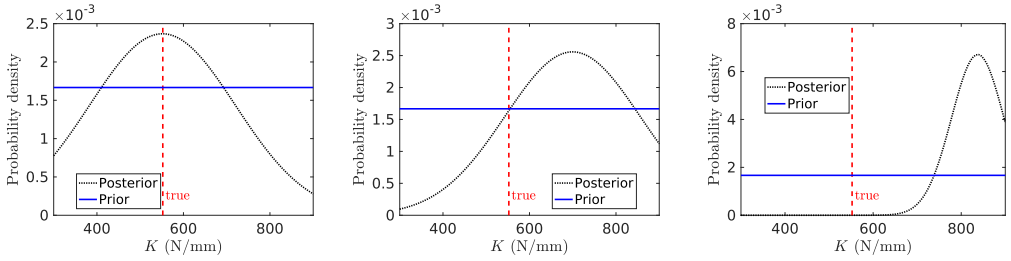


Figure 6.9: Posterior PDF of spring stiffness determined by the MBA, for minimum preposterior covariance values returned by the FPA, with a) three sensors b) four sensors and c) five sensors.

between the bias and the variance of the posterior PDF is visible. One possible way to decrease the overall posterior variance would be to consider differently natured responses, such as displacement, load in addition to strain. In other words to increase the informativeness supplied by the structural system. But this is beyond the scope of this work and it has already been demonstrated by Arendt et al.

Hence if the user is interested in better identifiability, he should select a configuration in this central region, with as few sensors as possible. If instead precision is more relevant, a configuration with more sensors, also in this central region, is recommended. Finally, a tradeoff solution would lie between these two extremes.

It is important to mention that these results were obtained on the basis of non-informative prior knowledge, which would imply that it is an over-conservative solution. Despite this, from a total number of nine sensors an optimal solution with five sensors would reduce costs to almost 50 %.

6.5 Summary

In this chapter, a measurement system design (MSD) methodology has been developed, in order to maximise the identifiability of st-id performed by the MBA, was presented in Chapter 5.

This methodology takes into account several sources of uncertainty, such as: parametric uncertainty, model discrepancy and observation error. If prior information regarding these uncertainties is available, e.g. knowledge regarding precision of a sensor/deficiencies of a model, it can be efficiently combined to obtain a more reliable sensor configuration. Alternatively non-informative uniform priors are used to obtain all these sources of uncertainty.

Subsequently, a Monte Carlo run propagates all these uncertainties and explores this input space, while recovering a preposterior covariance performance index. This criterion translates the variance of the posterior PDF, and is obtained by a FPA algorithm, detailed in the previous sections.

Highlights of this algorithm include a different numerical approach, which allows to reduce its computational effort, rendering it feasible to SHM applications, and an improved numerical efficiency, by exploiting the Cholesky factor decomposition of the correlation matrix, as described in Section 3.7.

Finally, this algorithm has been validated against a st-id combinatorial analysis of an aluminium bridge truss bridge case-study, already presented in Chapter 5. Additionally, a comparison of the computational effort required by this algorithm, against previous work from other authors was also detailed.

Chapter 7

Conclusions

In this chapter, major contributions and future work related with the uncertainty quantification of Structural Health Monitoring (SHM) based on the MBA are detailed.

In the introduction chapter a brief overview of the potential applications of the SHM technology, its economical and life-safety advantages was given. This information was complemented by some of the challenges that have to be surpassed for global acceptance of SHM, and in particular due to the nefarious presence of uncertainties. Finally, a description of all the uncertainties that compromise data-interpretation with a physics-based model was presented.

Chapter 2 highlighted the existent literature review on uncertainty quantification in SHM. A comprehensive description of several probabilistic methodologies has been detailed, with special emphasis on Bayesian probabilities, fuzzy numbers, Kalman filters, model falsification and sampling methods. Bayesian methodologies assumed a significant fraction of this body of research, having numerous applications for model updating, structural identification, measurement system design and operational modal analysis, been shown. It has been concluded that application of the SHM concept to complex infrastructure is very susceptible to uncertainties, and environmental/operational effects and model discrepancy severely compromise physics-based approaches. Finally, it was gathered that state of the art Bayesian frameworks which address these challenges present a number of shortcomings.

Review of past literature justified the development of an enhanced MBA, specifically tailored to address the issues mentioned in the previous chapter. The formulation of the MBA has been rigorously detailed in Chapter 4. In contrast with previous implementations of the MBA, the developed framework allows to identify multiple calibration parameters and is more stable, due to a Cholesky factorisation

procedure, described in section 3.7. A detailed range of tests with a simulated example has also been shown, illustrating the algorithm performance under different types of uncertainty, e.g. large model discrepancy, noise, outliers, diverse prior information, and multiple parameters identification. Results revealed that:

- large model discrepancy can severely affect the identifiability of the methodology, and it is recommended that several, diversified responses are used in order to improve the accuracy of the inference process;
- having stated the problematic issue of identifiability, when applied successfully, the methodology allows to uncover patterns and correlations related to the predictive ability of the model, which could be used for model updating;
- moreover, accurately identified parameters which reflect a performance of interest, e.g. stiffness or cross-sectional area of a structural component, can be used for damage detection. This last remark has not been tested but represents a potential future work;
- the MBA performs a reliable st-id, even in the presence of a low SNR and outliers, as seen in Section 4.7.4 and 4.7.6. Increasing the data set size further increases the precision of the st-id;
- the methodology scales well for multiple parameters inference, and the Monte Carlo numerical integration and MCMC sampling algorithms implemented in module 2 and 3 of the MBA have been successfully validated.

Chapter 5 illustrated the first practical applications of the MBA for st-id in SHM. Namely it has been applied to a reduced scale aluminium bridge and to the Tamar long suspension bridge. Based on results the following conclusions can be highlighted:

- the MBA is able to identify the true value of a structural parameter and predict responses, while considering uncertainties due to parametric variability, observation error, residual variability, interpolation uncertainty and model discrepancy for an experimental setup.
- it is possible to infer a structural parameter with only two measurements of a single characteristic nature, even in the presence of high discrepancies as shown by a comparative study. This could be further improved by adding additional responses (three, four) with a more diversified nature (inclination, displacement, etc.);

- identifiability is influenced by the dependency of structural parameters on measured responses. This is shown by measurements near singularities, such as supports of the bridge, which present poor identifiability relatively to measurements at the middle of the bridge. This is so because the dependency of the spring on the strain is smaller near supports;

- for the Tamar bridge example, a comparison of identified parameters was made against other authors, and the MBA indicates a more reasonable agreement with monitored data. Namely, two of the three identified parameters of the Tamar bridge have been validated with in-situ tests in Section 5.4.3;

- the MBA is also able to effectively highlight trends and patterns of model discrepancy, as seen in Section 5.4.2, which is an essential tool for analysts and decision-makers who have to rely on a model's predictive ability. Namely, in this study the traffic effects were modelled poorly, whereas the temperature only exhibited a localised discrepancy;

- since modelling temperature in this simplified fashion provided sufficient insight into the identified structural parameters, a transient thermal analysis does not seem to be strictly necessary. Provided with this information, a structural analyst can save a considerable amount of time and effort devoted to modelling a certain phenomenon.

In Chapter 6 a fast preposterior analysis has been developed, in order to select a sensor configuration which improves the identifiability of the MBA. No monitoring data is required for this analysis. Results of its application to a simulated cantilever beam example and a reduced-scale aluminium bridge have been validated against a st-id combinatorial analysis performed with the MBA. Main conclusions can be stated as follows:

- by using numerical strategies such as the Cholesky factor decomposition, and avoiding continuous estimation of mrGp hyperparameters in a MC loop of the original methodology, it is possible to reduce the computational effort to one third, of what was required by the original formulation.

- the results of Section 6.3.1 suggest that responses which have a high dependency on the structural parameters and model discrepancy provide the best structural identification;

- in Section 6.4 the same approach has been applied to the aluminium bridge case-study. Results are encouraging, despite the occurrence of a bias-variance

trade-off. Further tests are required, in order to improve this solution;

- in short, the FPA depends more strongly on the available prior knowledge than the MBA. Particularly on problems with high model discrepancy.

7.1 Achievements/contributions

Recalling the main objectives of this thesis, stated in Section 1.3, the following points constitute the core contributions from the current thesis

- Multiple parameter st-id extension (through numerical integration and implementation of the Metropolis-Hastings algorithm) and increase of the numerical stability (through the Cholesky factorisation) of the MBA.
- The implemented version of the MBA is capable to use more than two responses for st-id and MSD. At least eight responses have been tested for the Tamar bridge example.
- First practical application of the MBA for st-id in practical engineering applications, namely for a reduced scale model of an aluminium bridge and the Tamar long suspension bridge.
- Results of application of the FPA to a simulated example of a cantilever beam indicate (in a broad sense) the type of combination of responses which provide best identifiability.
- Finally, the first practical application of the FPA has also been tested in the aluminium bridge example.

The following scientific publications have been published/submitted on the basis of the work presented in this thesis

- Jesus, A., Brommer, P., Zhu, Y. AND Laory, I. Comprehensive Bayesian Structural Identification Using Temperature Variation. *Engineering Structures* 141 (June 2017), 75 –82. ISSN: 0141-0296. DOI: 10.1016/j.engstruct.2017.01.060.

- Jesus, A., Brommer, P., Westgate, R., Koo, K.-Y., Brownjohn, J. M. W. AND Laory, I. Bayesian Structural Identification of a Long Suspension Bridge Considering Temperature and Traffic Load Effects. *in review Journal of Structural Health Monitoring* (20th Mar. 2018).

- Jesus, A., Zhu, Y. AND Laory, I. Comprehensive Bayesian Structural Identification Using Temperature Expansion of a Scale Aluminium Bridge. In: *8th European Workshop on Structural Health Monitoring, EWSHM 2016*. Vol. 4. 2016, 3080–3088.

- Jesus, A., Salami, M., Westgate, R., Koo, K.-Y., Brownjohn, J. AND Laory, I. Bayesian Structural Identification of a Suspension Bridge Using Temperature and Traffic Loading. In: *Proceedings of the 8th International Conference on Structural Health Monitoring of Intelligent Infrastructure (SHMII), Brisbane*. 2017.

- Jesus, A., Zhu, Y., Koo, K.-Y., Brownjohn, J. AND Laory, I. Identifiability Based Sensor Configuration Analysis for Bayesian Structural Identification. In: *Proceedings of the 8th International Conference on Structural Health Monitoring of Intelligent Infrastructure (SHMII), Brisbane*. 2017.

7.2 Future work

Finally, future work related with the application of the MBA for Structural Health Monitoring is discussed in this section. Topics include model updating, operational modal analysis, damage detection and combination with a hierarchical Bayes model.

7.2.1 Model updating

As stated in the Introduction chapter, the methodology developed in this thesis was primarily designed to calibrate and validate a structural model against monitored data, while quantifying its associated uncertainties.

Therefore, it is relevant to include it as a decision-making module for a model updating formulation. An example of such a framework can be seen in Section 4, Fig. 2.8, where calibration parameters are updated continuously as more and more data is being supplied to the algorithm. By the same token, it is also possible to update tuning parameters e.g. FE mesh size, to obtain a more refined and accurate model. This would essentially require establishment of a predicitive performance criteria, such as model discrepancy below a certain threshold, or variability of identified model parameters.

7.2.2 Operational modal analysis

Regarding the applicability of the MBA for operational modal analysis it would be interesting to consider a forced damped vibration decay model as suggested by

Au [82] for estimation of natural frequencies, mode shapes and damping coefficients.

Suggestion of this future work is justified by the fact that, results of Au Bayesian approach tended to be potentially biased, specially for modal identification during strong winds. Naturally the effects of wind, waves and other natural loading excitation, could be represented as a discrepancy function with an associated stochastic process [217]. It would be interesting to see if the multiple response Gaussian process of the MBA would be able to represent this bias, and improve the identification process.

However, it should be noted that a large amount of data at a high sampling rate is usually necessary to perform this type of analyses, and although the MBA uses mrGp matrices in sparse format, in the presence of such a large amount of data, these could easily become ill-conditioned.

Possible suggestions to overcome this hurdle include applying a model updating process as suggested in the previous section, where the only information saved from step to step are the key features of the calibration process, i.e. posterior of updated parameters, mrGp hyperparameters, etc., or alternatively to use a different type of surrogate model, that better handles large amounts of data, such as the deep Gaussian processes from Damianou et al. [218, 219].

7.2.3 Damage detection

Another very promising area of research for the MBA is damage detection in civil infrastructure. The type of damage detection could be either supervised or unsupervised learning. In the supervised learning approach, the structural model would firstly have to be calibrated with relevant environmental and operational conditions.

For example, assuming that it is possible to establish a reference state, when the structural system is assumed healthy and its parameters are estimated, a subsequent comparison of this reference with an estimate of the parameters for a current state would indicate the presence, location and extent of damage. A flowchart of such algorithm can be seen in Fig. 7.1, where a damage detection framework requires computation of a mrGp of the computer model identically to the original MBA, in Task 1. Subsequently the discrepancy function and calibration parameters posterior are estimated in a reference and current state, Task 2 and 3, respectively. Each of the tasks iterates over module 2 and 3 from the MBA, supplying prior information of the calibration parameter and monitored data D_r^e or D_c^e. During such computations it would be advisable to supply a monitored reference dataset D_r^e which includes as much information as possible, e.g. seasonal variations in a one year time frame, and is acquired at the earliest possible stage of the structure' life-cycle. Finally, in Task

4 the uncertainties of the parameters posteriors are propagated to a damage metric random variable DF.

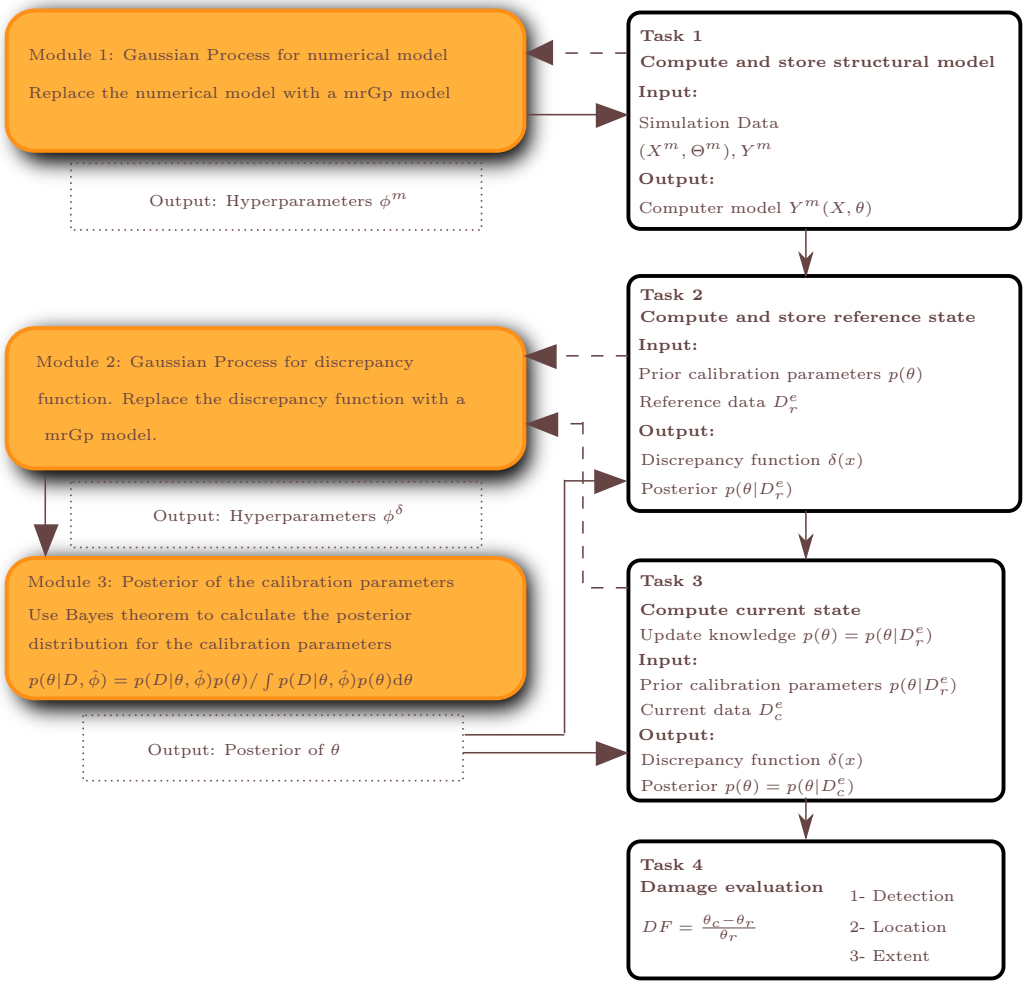


Figure 7.1: Flowchart of the MBA original approach (left) and the proposed damage detection framework (right).

On the other hand, in an unsupervised learning approach a damage index related with the occurrence of damage could be based on the variability of the model discrepancy. Advantages of this framework include its global ability to detect the presence of damage, although in this scenario the predictive ability of the model to locate, measure the extension and future progression of damage could not be assessed.

Obviously it would be necessary to correctly establish the balance between Type I and Type II errors of such a methodology, see Sec. 6.1 for more details. Nevertheless, the fact that the MBA is such a comprehensive methodology, and

that Gaussian processes are designed for tasks such as regression and classification, suggests it as a promising candidate for damage detection.

7.2.4 Inclusion of hierarchical Bayes model

Finally, it has been stated extensively in the literature review Chapter 2 of this thesis, that compared to the hierarchical Bayes framework the MBA does not take into account the inherent variability of model parameters. In other words, as the amount of calibration data converges to infinity the PDF of the identified parameters converges to a Dirac delta function, and not to a PDF which translates the variability of the parameter during the monitoring process. Such detail is very relevant for SHM applications, where structural parameters change during the life-time of an infrastructure, because it allows to separate influences due to damage from variability of estimated parameters.

One possible solution to extend the MBA for such type of problem would be to consider the actual moments (mean, covariance) of a target distribution as parameters to identify, and run a much larger set of simulations with different moments. This is not very efficient, since it would require an extremely large number of simulations to fit the model mrGp and it would increase much more quickly the number of parameters to identify.

The more elegant solution would be to develop a methodology able to quantify qualitatively and quantitatively all sources of uncertainty mentioned in Section 2.2.2, under a hierarchical Bayes model. The combination of the hierarchical and modular frameworks could be established on the basis of some hierarchical Gaussian process, as presented by Schwaighofer et al. [220]. Despite the inherent difficulties in order to implement and test such methodology, the author believes it would bring considerable improvements to the usage and global acceptance of the SHM concept.

Appendix A

Conditional distribution of multivariate normal

Assuming a multivariate normal vector $\boldsymbol{Y} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$. Consider partitioning $\boldsymbol{\mu}$ and $\boldsymbol{Y}$ into

$$\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}$$

$$\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \end{bmatrix}$$

with a similar partition of Σ into

$$\begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

Then,$(\boldsymbol{y}_1 | \boldsymbol{y}_2 = \boldsymbol{a})$, the conditional distribution of the first partition given the second, is, $\mathcal{N}(\overline{\boldsymbol{\mu}}, \overline{\Sigma})$ with mean

$$\overline{\boldsymbol{\mu}} = \boldsymbol{\mu}_1 + \Sigma_{12}{\Sigma_{22}}^{-1}(\boldsymbol{a} - \boldsymbol{\mu}_2)$$

and covariance matrix

$$\overline{\Sigma} = \Sigma_{11} - \Sigma_{12}{\Sigma_{22}}^{-1}\Sigma_{21}$$

A.1 Proof

Let $\mathbf{x_1}$ be the first partition and $\mathbf{x_2}$ the second. Now define $\mathbf{z} = \mathbf{x_1} + \mathbf{A}\mathbf{x_2}$ where $\mathbf{A} = -\Sigma_{12}\Sigma_{22}^{-1}$. Now it is possible to write

$$\begin{aligned}
\text{cov}(\mathbf{z}, \mathbf{x}_2) &= \text{cov}(\mathbf{x}_1, \mathbf{x}_2) + \text{cov}(\mathbf{A}\mathbf{x}_2, \mathbf{x}_2) \\
&= \Sigma_{12} + \mathbf{A}\text{var}(\mathbf{x}_2) \\
&= \Sigma_{12} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{22} \\
&= 0
\end{aligned}$$

Therefore $\mathbf{z}$ and $\mathbf{x_2}$ are uncorrelated and, since they are jointly normal, they are independent. Now, clearly $E(\mathbf{z}) = \boldsymbol{\mu}_1 + \mathbf{A}\boldsymbol{\mu}_2$, therefore it follows that

$$\begin{aligned}
E(\mathbf{x}_1|\mathbf{x}_2) &= E(\mathbf{z} - \mathbf{A}\mathbf{x}_2|\mathbf{x}_2) \\
&= E(\mathbf{z}|\mathbf{x}_2) - E(\mathbf{A}\mathbf{x}_2|\mathbf{x}_2) \\
&= E(\mathbf{z}) - \mathbf{A}\mathbf{x}_2 \\
&= \boldsymbol{\mu}_1 + \mathbf{A}(\boldsymbol{\mu}_2 - \mathbf{x}_2) \\
&= \boldsymbol{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)
\end{aligned}$$

which proves the first part. For the covariance matrix, note that

$$\begin{aligned}
\text{var}(\mathbf{x}_1|\mathbf{x}_2) &= \text{var}(\mathbf{z} - \mathbf{A}\mathbf{x}_2|\mathbf{x}_2) \\
&= \text{var}(\mathbf{z}|\mathbf{x}_2) + \text{var}(\mathbf{A}\mathbf{x}_2|\mathbf{x}_2) - \mathbf{A}\text{cov}(\mathbf{z}, -\mathbf{x}_2) - \text{cov}(\mathbf{z}, -\mathbf{x}_2)\mathbf{A}' \\
&= \text{var}(\mathbf{z}|\mathbf{x}_2) \\
&= \text{var}(\mathbf{z})
\end{aligned}$$

Finally

$$\begin{aligned}
\text{var}(\mathbf{x}_1|\mathbf{x}_2) = \text{var}(\mathbf{z}) &= \text{var}(\mathbf{x}_1 + \mathbf{A}\mathbf{x}_2) \\
&= \text{var}(\mathbf{x}_1) + \mathbf{A}\text{var}(\mathbf{x}_2)\mathbf{A}' + \mathbf{A}\text{cov}(\mathbf{x}_1, \mathbf{x}_2) + \text{cov}(\mathbf{x}_2, \mathbf{x}_1)\mathbf{A}' \\
&= \Sigma_{11} + \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{22}\Sigma_{22}^{-1}\Sigma_{21} - 2\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} \\
&= \Sigma_{11} + \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} - 2\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} \\
&= \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}
\end{aligned}$$

which proves the second part.

Appendix B

Estimating posterior distribution with regression coefficients

This appendix details the marginalisation of the posterior distribution

$$p(\boldsymbol{\theta}, \boldsymbol{\phi}, \boldsymbol{\beta} | \boldsymbol{Y}, \boldsymbol{D}) \propto p(\boldsymbol{\theta}) p(\boldsymbol{\phi}) |\boldsymbol{V}_Y(\boldsymbol{\theta})|^{-1/2} \times \exp\left\{ -\frac{1}{2} (\boldsymbol{Y} - \boldsymbol{H}_Y(\boldsymbol{\theta})\boldsymbol{\beta})^T \boldsymbol{V}_Y(\boldsymbol{\theta})^{-1} (\boldsymbol{Y} - \boldsymbol{H}_Y(\boldsymbol{\theta})\boldsymbol{\beta}) \right\}, \quad \text{(B.1)}$$

in order to the regression coefficients $\boldsymbol{\beta}$, which have been estimated in section 3.4.1 through MLE. In Eq. (B.1) $\boldsymbol{\theta}$, $\boldsymbol{\phi}$, $\boldsymbol{Y}$, $\boldsymbol{D}$ represent calibration parameters, hyperparameters of mrGps and sets of output and input-output data, respectively. Lastly, the regression matrix $\boldsymbol{H}_Y$ and covariance matrix $\boldsymbol{V}_Y$ define the mrGp which regresses the data, and allow to infer $\boldsymbol{\theta}$ (note the dependency in the above equation).

The objective is to demonstrate that such marginalisation results in the following expression

$$p(\boldsymbol{\theta} | \boldsymbol{Y}, \boldsymbol{D}, \hat{\boldsymbol{\phi}}) \propto p(\boldsymbol{\theta}) p(\boldsymbol{\phi}) |\boldsymbol{V}_Y(\boldsymbol{\theta})|^{-1/2} |\boldsymbol{W}_Y(\boldsymbol{\theta})|^{1/2} \times \exp\left\{ -\frac{1}{2} (\boldsymbol{Y} - \boldsymbol{H}_Y(\boldsymbol{\theta})\hat{\boldsymbol{\beta}})^T \boldsymbol{V}_Y(\boldsymbol{\theta})^{-1} (\boldsymbol{Y} - \boldsymbol{H}_Y(\boldsymbol{\theta})\hat{\boldsymbol{\beta}}) \right\}, \quad \text{(B.2)}$$

where $\boldsymbol{W}_Y$ is the Gramian matrix, already discussed in Section 3.4.1.

B.1 Proof

As detailed previously the MLE of the regression coefficients is

$$\begin{aligned} \hat{\boldsymbol{\beta}}(\boldsymbol{\theta}) &= \boldsymbol{W}_Y(\boldsymbol{\theta})\boldsymbol{H}_Y(\boldsymbol{\theta})^T\boldsymbol{V}_Y(\boldsymbol{\theta})^{-1}\boldsymbol{Y} \\ \boldsymbol{W}_Y(\boldsymbol{\theta}) &= (\boldsymbol{H}_Y(\boldsymbol{\theta})^T\boldsymbol{V}_Y(\boldsymbol{\theta})^{-1}\boldsymbol{H}_Y(\boldsymbol{\theta}))^{-1} \end{aligned} \tag{B.3}$$

due to the assumption of normality of the coefficients

$$\boldsymbol{\beta}|\boldsymbol{\theta},\boldsymbol{\phi},\boldsymbol{Y} \sim \mathcal{N}(\hat{\boldsymbol{\beta}}(\boldsymbol{\theta}), \boldsymbol{W}_Y(\boldsymbol{\theta})) \tag{B.4}$$

For the sake of clarity, the calibration parameters $\boldsymbol{\theta}$ will be omitted in the subsequent deductions from $\boldsymbol{W}_Y, \boldsymbol{H}_Y, \boldsymbol{V}_Y$ and $\boldsymbol{\beta}$. Firstly, it is necessary to expand the quadratic form in the exponent of Eq. (B.1) as follows

$$\begin{aligned} p(\boldsymbol{\theta},\boldsymbol{\beta},\boldsymbol{\phi}|\boldsymbol{Y}) \propto\ & p(\boldsymbol{\theta})p(\boldsymbol{\phi})|\boldsymbol{V}_Y|^{-1/2}\exp\Bigg\{-\frac{1}{2}\Bigg[\boldsymbol{Y}^T\boldsymbol{V}_Y^{-1}\boldsymbol{Y} \\ & -\boldsymbol{\beta}^T\overbrace{\boldsymbol{H}_Y^T\boldsymbol{V}_Y^{-1}\boldsymbol{Y}}^{\boldsymbol{W}_Y^{-1}\hat{\boldsymbol{\beta}}} - \overbrace{\boldsymbol{Y}^T\boldsymbol{V}_Y^{-1}\boldsymbol{H}_Y}^{\hat{\boldsymbol{\beta}}^T\boldsymbol{W}_Y^{-1}}\boldsymbol{\beta} + \\ & +\boldsymbol{\beta}^T\overbrace{\boldsymbol{H}_Y^T\boldsymbol{V}_Y^{-1}\boldsymbol{H}_Y}^{\boldsymbol{W}_Y^{-1}}\boldsymbol{\beta}\Bigg]\Bigg\}. \end{aligned}$$

which by replacing the MLE noted in the overbraces, results in

$$\begin{aligned} p(\boldsymbol{\theta},\boldsymbol{\beta},\boldsymbol{\phi}|\boldsymbol{Y}) \propto\ & p(\boldsymbol{\theta})p(\boldsymbol{\phi})|\boldsymbol{V}_Y|^{-1/2}\exp\left\{-\frac{1}{2}[\boldsymbol{Y}^T\boldsymbol{V}_Y^{-1}\boldsymbol{Y}]\right\} \\ & \times\exp\left\{-\frac{1}{2}[\boldsymbol{\beta}^T\boldsymbol{W}_Y(\boldsymbol{\theta})^{-1}\boldsymbol{\beta} - 2\boldsymbol{\beta}^T\boldsymbol{W}_Y^{-1}\hat{\boldsymbol{\beta}}]\right\}. \end{aligned}$$

It is now necessary to add and subtract $\hat{\boldsymbol{\beta}}^T\boldsymbol{W}_Y^{-1}\hat{\boldsymbol{\beta}}$ to the second exponent, which then develops into

$$\begin{aligned} & p(\boldsymbol{\theta},\boldsymbol{\beta},\boldsymbol{\phi}|\boldsymbol{Y}) \propto p(\boldsymbol{\theta})p(\boldsymbol{\phi})|\boldsymbol{V}_Y|^{-1/2}\exp\left\{-\frac{1}{2}[\boldsymbol{Y}^T\boldsymbol{V}_Y^{-1}\boldsymbol{Y}]\right\} \\ & \times\exp\left\{-\frac{1}{2}[\boldsymbol{\beta}^T\boldsymbol{W}_Y^{-1}\boldsymbol{\beta} - 2\boldsymbol{\beta}^T\boldsymbol{W}_Y^{-1}\hat{\boldsymbol{\beta}} + \hat{\boldsymbol{\beta}}^T\boldsymbol{W}_Y^{-1}\hat{\boldsymbol{\beta}}]\right\} \\ & \times\exp\left\{+\frac{1}{2}[\hat{\boldsymbol{\beta}}^T\boldsymbol{W}_Y^{-1}\hat{\boldsymbol{\beta}}]\right\}. \end{aligned} \tag{B.5}$$

At this point it can be observed that the second exponent is the quadratic expansion of the regression coefficients $\boldsymbol{\beta}$ normal equation, with mean $\hat{\boldsymbol{\beta}}$ and covariance $\boldsymbol{W}_Y$.

Thus, Eq. (B.5) can be written in a simplified form

$$\begin{aligned} p(\boldsymbol{\theta},\boldsymbol{\beta},\boldsymbol{\phi}|\boldsymbol{Y}) &\propto p(\boldsymbol{\theta})p(\boldsymbol{\phi})|\boldsymbol{V}_Y|^{-1/2}\exp\left\{-\frac{1}{2}[\boldsymbol{Y}^T\boldsymbol{V}_Y^{-1}\boldsymbol{Y}]\right\} \\ &\quad\times\exp\left\{-\frac{1}{2}[(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})^T\boldsymbol{W}_Y^{-1}(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})+\frac{1}{2}[\hat{\boldsymbol{\beta}}^T\boldsymbol{W}_Y^{-1}\hat{\boldsymbol{\beta}}]\right\} \end{aligned}$$

Since Eq. (B.6) only depends of the regression coefficients on its second term, and that term is a normal equation, its integration in order to $\boldsymbol{\beta}$ will yield the square root of the determinant of the Gramian matrix. The other terms remain unchanged.

$$\begin{aligned} \int p(\boldsymbol{\theta},\boldsymbol{\beta},\boldsymbol{\phi}|\boldsymbol{Y})\mathrm{d}\boldsymbol{\beta} &\propto p(\boldsymbol{\theta})p(\boldsymbol{\phi})|\boldsymbol{V}_Y|^{-1/2}\exp\left\{-\frac{1}{2}[\boldsymbol{Y}^T\boldsymbol{V}_Y^{-1}\boldsymbol{Y}]\right\} \\ &\quad|\boldsymbol{W}_Y|^{1/2}\exp\left\{+\frac{1}{2}[\hat{\boldsymbol{\beta}}^T\boldsymbol{W}_Y^{-1}\hat{\boldsymbol{\beta}}]\right\} \end{aligned} \tag{B.6}$$

which can be reordered as

$$\begin{aligned} p(\boldsymbol{\theta},\boldsymbol{\phi}|\boldsymbol{Y}) &\propto p(\boldsymbol{\theta})p(\boldsymbol{\phi})|\boldsymbol{V}_Y|^{-1/2}|\boldsymbol{W}_Y|^{1/2} \\ &\quad\exp\left\{-\frac{1}{2}[\boldsymbol{Y}^T\boldsymbol{V}_Y^{-1}\boldsymbol{Y}-\hat{\boldsymbol{\beta}}^T\boldsymbol{W}_Y^{-1}\hat{\boldsymbol{\beta}}]\right\} \end{aligned}$$

The final task is now to simplify the remaining exponent term. The first requirement is to sum and subtract $2\hat{\boldsymbol{\beta}}^T\boldsymbol{W}_Y^{-1}\hat{\boldsymbol{\beta}}$ on the exponent

$$\begin{aligned} &p(\boldsymbol{\theta},\boldsymbol{\phi}|\boldsymbol{Y})\propto p(\boldsymbol{\theta})p(\boldsymbol{\phi})|\boldsymbol{V}_Y|^{-1/2}|\boldsymbol{W}_Y|^{1/2} \\ &\exp\left\{-\frac{1}{2}[\boldsymbol{Y}^T\boldsymbol{V}_Y^{-1}\boldsymbol{Y}-2\hat{\boldsymbol{\beta}}^T\boldsymbol{W}_Y^{-1}\hat{\boldsymbol{\beta}}+\hat{\boldsymbol{\beta}}^T\boldsymbol{W}_Y^{-1}\hat{\boldsymbol{\beta}}]\right\} \end{aligned}$$

and by carefully replacing the MLE of the regression coefficients in the second term and the Gramian matrix inverse in the third term yields

$$\begin{aligned} p(\boldsymbol{\theta},\boldsymbol{\phi}|\boldsymbol{Y}) &\propto p(\boldsymbol{\theta})p(\boldsymbol{\phi})|\boldsymbol{V}_Y|^{-1/2}|\boldsymbol{W}_Y|^{1/2}\exp\Big\{-\frac{1}{2}[\boldsymbol{Y}^T\boldsymbol{V}_Y^{-1}\boldsymbol{Y} \\ &\quad -2\hat{\boldsymbol{\beta}}^T\boldsymbol{W}_Y^{-1}\overbrace{\boldsymbol{W}_Y\boldsymbol{H}_Y^T\boldsymbol{V}_Y^{-1}\boldsymbol{Y}}^{\hat{\boldsymbol{\beta}}}+\hat{\boldsymbol{\beta}}^T\overbrace{\boldsymbol{H}_Y^T\boldsymbol{V}_Y^{-1}\boldsymbol{H}_Y}^{\boldsymbol{W}_Y^{-1}}\hat{\boldsymbol{\beta}}]\Big\} \end{aligned}$$

Finally, note that the Gramian matrix cancels out with its inverse in the second term inside of the exponent. The final equation is, once again, a normal equation of mean

$\boldsymbol{H}_Y\hat{\boldsymbol{\beta}}$ and covariance $\boldsymbol{V}_Y$

$$\begin{aligned} p(\boldsymbol{\theta},\boldsymbol{\phi}|\boldsymbol{Y}) &\propto p(\boldsymbol{\theta})p(\boldsymbol{\phi})|\boldsymbol{V}_Y|^{-1/2}|\boldsymbol{W}_Y|^{1/2} \\ &\quad \exp\left\{-\frac{1}{2}[\boldsymbol{Y}^T\boldsymbol{V}_Y^{-1}\boldsymbol{Y} - 2\hat{\boldsymbol{\beta}}^T\boldsymbol{H}_Y^T\boldsymbol{V}_Y^{-1}\boldsymbol{Y} + \hat{\boldsymbol{\beta}}^T\boldsymbol{H}_Y^T\boldsymbol{V}_Y^{-1}\boldsymbol{H}_Y\hat{\boldsymbol{\beta}}]\right\} \end{aligned}$$

which finally results in

$$\begin{aligned} p(\boldsymbol{\theta},\boldsymbol{\phi}|\boldsymbol{Y}) &\propto p(\boldsymbol{\theta})p(\boldsymbol{\phi})|\boldsymbol{V}_Y|^{-1/2}|\boldsymbol{W}_Y|^{1/2} \\ &\quad \times \exp\left\{-\frac{1}{2}(\boldsymbol{Y}-\boldsymbol{H}_Y\hat{\boldsymbol{\beta}})^T\boldsymbol{V}_Y^{-1}(\boldsymbol{Y}-\boldsymbol{H}_Y\hat{\boldsymbol{\beta}})\right\} \end{aligned} \tag{B.7}$$

which concludes the proof of Eq. (4.30).

Appendix C

Numerical approach

Solving directly the modular Bayesian formulation is not practical from a numerical point of view. Several numerical considerations are required, in order to establish a robust and light-weight algorithm.

Some examples that need special consideration include

1. demanding operations such as the inverse of the covariance/correlation matrices in Eqs. (4.5), (4.18) and (4.30), which can easily become large and ill-conditioned;

2. inversion of these same matrices is computationally demanding and can lead to stability problems;

3. likewise, the stability of the likelihood functions which are optimised in modules 1, 2, and 3, should be replaced by their more tractable log-likelihood equivalents;

4. since the algorithm has a modular structure, it can be applied to a wide range of applications, and it is convenient to work with standardised units (dimensionless quantities) rather than with absolute values. This increases numerical efficiency;

Some numerical techniques that aid to solve this issues include: the Cholesky/ QR decomposition, the Kronecker product, formatting of matrices in sparse format, and the standardisation of the datasets. Relevant equations, which are directly resultant from these simplifications will be rigorously detailed and highlighted.

C.1 Cholesky and QR decomposition

In this section, the Cholesky and QR decomposition factorisation will be used jointly, in order to solve issues 1 and 2. This factorisation is a standard procedure when dealing with a generalised least squares problem [221], and it basically consists of an orthogonal projection, which reduces the generalised least squares problem, to the solution of the ordinary least squares problem.

Application of the Cholesky factor decomposition is straightforward since the correlation or covariance matrices in this thesis are symmetric and positive-definite. These two properties imply that they can be written in factorised form

$$\boldsymbol{R} = \boldsymbol{C}\boldsymbol{C}^T, \tag{C.1}$$

with $\boldsymbol{C}$ a lower triangular matrix. This transformation allows to perform an orthogonal transformation from the generalised to ordinary least squares solution of the overdetermined system:

$$\widetilde{\boldsymbol{H}}\boldsymbol{\beta} \simeq \widetilde{\boldsymbol{Y}} \tag{C.2}$$

with

$$\boldsymbol{C}\widetilde{\boldsymbol{H}} = \boldsymbol{H}, \qquad \boldsymbol{C}\widetilde{\boldsymbol{Y}} = \boldsymbol{Y}. \tag{C.3}$$

additionally the "economy size" QR factorization of $\widetilde{\boldsymbol{H}}$ is

$$\widetilde{\boldsymbol{H}} = \boldsymbol{Q}\boldsymbol{G}^T \tag{C.4}$$

where $\boldsymbol{Q}$ is an orthogonal matrix and $\boldsymbol{G}^T$ is an upper triangular matrix. Having presented the two factorisations, it is now very simple to replace the transformed dataset $\widetilde{\boldsymbol{Y}}$, regression function $\widetilde{\boldsymbol{H}}$ and the covariance matrix of the model to simplify the likelihood functions shown in the original formulation.

C.1.1 Properties of Kronecker product and Cholesky/QR decomposition

A few of the properties that are helpful to understand the simplified likelihood functions of the MBA require knowledge of a few properties, related with how the Cholesky and QR factorisation do operate with the Kronecker product. These shall be stated in the following sections.

Determinant

The covariance matrix of the model, in Eq. (4.6) for example, can be written as

$$\boldsymbol{V}_m = \boldsymbol{\Sigma}_m^2 \otimes \boldsymbol{C}_m \boldsymbol{C}_m^T \tag{C.5}$$

and the determinant of this matrix can be written as

$$|\boldsymbol{V}_m| = |\boldsymbol{\Sigma}_m^2 \otimes \boldsymbol{R}_m| = |\boldsymbol{\Sigma}_m^2|^N |\boldsymbol{R}_m|^q = |\boldsymbol{\Sigma}_m^2|^N |\boldsymbol{C}_m|^{2q} \tag{C.6}$$

since $|\boldsymbol{R}| = |\boldsymbol{C}|^2$, where N and q are the size of correlation matrix $\boldsymbol{R}$, and of the variance matrix $\boldsymbol{\Sigma}^2$, respectively.

Inverse

Another useful property of the Kronecker product is the distributive of the inverse of the product of two matrices, i.e.

$$(\boldsymbol{\Sigma}_m^2 \otimes \boldsymbol{R}_m)^{-1} = (\Sigma_m^2)^{-1} \otimes \ \boldsymbol{R}_m^{-1} \tag{C.7}$$

Mixed product property

If necessary the Kronecker product also has a mixed product property that states

$$(\boldsymbol{I}_q \boldsymbol{\Sigma}_m^2) \otimes (\boldsymbol{C}_m \boldsymbol{C}_m^T) = (\boldsymbol{I}_q \otimes \boldsymbol{C}_m)(\boldsymbol{\Sigma}_m^2 \otimes \boldsymbol{C}_m^T), \tag{C.8}$$

where $\boldsymbol{I}_q$ is the identity matrix with dimension q.

Equivalency property

Finally, a last equivalency property which is very useful to connect the Cholesky factorisation with the Kronecker product can be reviewed in Schäcke [196] and can be stated as follows. Assuming that there are matrices $\boldsymbol{V}, \boldsymbol{\Sigma}^2$, and $\boldsymbol{R}$, so that

$$\boldsymbol{V} = \boldsymbol{\Sigma}^2 \otimes \boldsymbol{R}, \tag{C.9}$$

with $\boldsymbol{\Sigma}^2$ and $\boldsymbol{R}$ symmetric then the following property applies

$$\boldsymbol{C}_V = \boldsymbol{C}_\Sigma \otimes \boldsymbol{C}_R, \tag{C.10}$$

with $\boldsymbol{C}_V, \boldsymbol{C}_\Sigma$ and $\boldsymbol{C}_R$ being the Cholesky decomposition matrices of $\boldsymbol{V}, \boldsymbol{\Sigma}^2$ and $\boldsymbol{R}$, respectively.

C.1.2 Module 1

Given the above stated transformations and their properties, it is now possible to apply these techniques in the log-likelihood function of module 1, in Eq. (4.8), Section 4.5.2. By applying the determinant expression of Eq. (C.6) and the transformed regression function and correlation matrix of Eq. (C.4) and Eq. (C.1) results in

$$\ell(y|\boldsymbol{\omega}_m, \boldsymbol{D}^m) \propto -\frac{1}{2}\{\log|\boldsymbol{\Sigma}_m^2|^N|\boldsymbol{C}_m|^{2q} - \log|((\boldsymbol{C}_m\boldsymbol{Q}_m\boldsymbol{G}_m^T)^T\boldsymbol{C}_m^{-T}\boldsymbol{C}_m^{-1}(\boldsymbol{C}_m\boldsymbol{Q}_m\boldsymbol{G}_m^T))^{-1}|\}, \tag{C.11}$$

which simplifies into

$$\ell(y|\boldsymbol{\omega}_m, \boldsymbol{D}^m) \propto -\frac{1}{2}\{\log|\boldsymbol{\Sigma}_m^2|^N|\boldsymbol{C}_m|^{2q} - \log|(\boldsymbol{G}_m\boldsymbol{G}_m^T)^{-1}|\}, \tag{C.12}$$

since most terms of the second term cancel out. Given this new log-likelihood function, numerically speaking, maximising this equation in order to the hyperparameters ω_m will be a much more efficient task.

C.1.3 Module 2

Similarly for Eq. (4.19) of the log-likelihood function in module 2 of the MBA, the same reasoning as the previous section can be applied, except that here, the matrix that is being factorized is the covariance matrix $\boldsymbol{V}_e$.

For simplicity of notation a discrepancy function dataset shall be designated as $\boldsymbol{Y}^\delta = \boldsymbol{Y}^e - \hat{\boldsymbol{Y}}^m$. Hence, this dataset and the regression functions $\boldsymbol{H}_\delta$ transformed forms will be

$$\boldsymbol{C}_e\widetilde{\boldsymbol{Y}}^\delta = \text{vec}(\boldsymbol{Y}^\delta) \qquad \boldsymbol{C}_e\widetilde{\boldsymbol{H}}_\delta = \boldsymbol{I}_q \otimes \boldsymbol{H}_\delta \tag{C.13}$$

similarly to what has been presented in Eq. (C.3)

The simplified log-likelihood function will therefore be expressed as

$$\begin{aligned} \ell(\boldsymbol{Y}^e|\boldsymbol{D}^m, \hat{\boldsymbol{\phi}}_m, \boldsymbol{\phi}_\delta) \quad &\propto \quad -\frac{1}{2}\{\log|\boldsymbol{C}_e|^2 - \log|(\boldsymbol{G}_e\boldsymbol{G}_e^T)^{-1}| \\ &+ \quad (\boldsymbol{C}^e\widetilde{\boldsymbol{Y}}^\delta - \boldsymbol{C}^e\widetilde{\boldsymbol{H}}_\delta\hat{\boldsymbol{\beta}}_\delta)^T\boldsymbol{C}_e^{-T}\boldsymbol{C}_e^{-1}(\boldsymbol{C}^e\widetilde{\boldsymbol{Y}}^\delta - \boldsymbol{C}^e\widetilde{\boldsymbol{H}}_\delta\hat{\boldsymbol{\beta}}_\delta)\}, \end{aligned} \tag{C.14}$$

which simplifies into

$$\ell(\boldsymbol{Y}^e|\boldsymbol{D}^m, \hat{\boldsymbol{\phi}}_m, \boldsymbol{\phi}_\delta) \propto -\frac{1}{2}\{\log|\boldsymbol{C}_e|^2 - \log|(\boldsymbol{G}_\delta\boldsymbol{G}_\delta^T)^{-1}| + (\widetilde{\boldsymbol{Y}}^\delta - \widetilde{\boldsymbol{H}}_\delta\hat{\boldsymbol{\beta}}_\delta)^T(\widetilde{\boldsymbol{Y}}^\delta - \widetilde{\boldsymbol{H}}_\delta\hat{\boldsymbol{\beta}}_\delta)\}. \tag{C.15}$$

In a similar fashion the regression coefficients $\boldsymbol{\beta}_\delta$ can also be efficiently computed as

$$\hat{\boldsymbol{\beta}}_\delta = \boldsymbol{G}_\delta^{-T} \boldsymbol{Q}_\delta^T \widetilde{\boldsymbol{Y}}^\delta, \tag{C.16}$$

which derives from Eqs. (4.21), (C.13), (C.4) and the discrepancy dataset.

These equations are considerably more simple than Eq. (4.19), and can be used to efficiently estimate the discrepancy function mrGp hyperparameters $\boldsymbol{\phi}^\delta$ and the observation error with genetic algorithms.

Specifically, an initial population of size 60 was generated in the [0–1] range, with default values for Gaussian mutation function (mean 0, standard deviation 1 and shrinkage of the standard deviation as generations go by 1) and scattered crossover function (0.8 fraction of the population at the each next generation). Convergence criteria are set as either a maximum number of 500 generations or an average change in the fitness value less than 1×10^{-3}. Since the algorithm often normalises the input data, these values have been used consistently across most of the tests presented in this thesis.

C.1.4 Module 3

Finally, Eq. (4.30) of module 3 presents a similar structure as the one displayed in the previous sections. By applying the same transformations on the global dataset, covariance matrix and regression functions this equation can be easily transformed into

$$\begin{aligned} \ell(\boldsymbol{Y}|\boldsymbol{\theta}, \boldsymbol{D}, \hat{\boldsymbol{\phi}}) \quad \propto \quad & -\frac{1}{2}\{\log|\boldsymbol{C}_Y(\boldsymbol{\theta})|^2 - \log|(\boldsymbol{G}_Y(\boldsymbol{\theta})\boldsymbol{G}_Y(\boldsymbol{\theta})^T)^{-1}| \\ + \quad & (\boldsymbol{C}_Y(\boldsymbol{\theta})\widetilde{\boldsymbol{Y}} - \boldsymbol{C}_Y(\boldsymbol{\theta})\widetilde{\boldsymbol{H}}_Y(\boldsymbol{\theta})\hat{\boldsymbol{\beta}})^T \boldsymbol{C}_Y(\boldsymbol{\theta})^{-T} \boldsymbol{C}_Y(\boldsymbol{\theta})^{-1} \\ & (\boldsymbol{C}_Y(\boldsymbol{\theta})\widetilde{\boldsymbol{Y}} - \boldsymbol{C}_Y(\boldsymbol{\theta})\widetilde{\boldsymbol{H}}_Y(\boldsymbol{\theta})\hat{\boldsymbol{\beta}})\}. \end{aligned} \tag{C.17}$$

which simplifies into

$$\begin{aligned} \ell(\boldsymbol{Y}|\boldsymbol{\theta}, \boldsymbol{D}, \hat{\boldsymbol{\phi}}) \quad \propto \quad & -\frac{1}{2}\{\log|\boldsymbol{C}_Y(\boldsymbol{\theta})|^2 - \log|(\boldsymbol{G}_Y(\boldsymbol{\theta})\boldsymbol{G}_Y(\boldsymbol{\theta})^T)^{-1}| \\ + \quad & (\widetilde{\boldsymbol{Y}} - \widetilde{\boldsymbol{H}}_Y(\boldsymbol{\theta})\hat{\boldsymbol{\beta}})^T (\widetilde{\boldsymbol{Y}} - \widetilde{\boldsymbol{H}}_Y(\boldsymbol{\theta})\hat{\boldsymbol{\beta}})\}. \end{aligned} \tag{C.18}$$

with

$$\begin{aligned} \hat{\boldsymbol{\beta}}(\boldsymbol{\theta}) &= \boldsymbol{G}_Y(\boldsymbol{\theta})^{-T} \boldsymbol{Q}_Y(\boldsymbol{\theta})^T \widetilde{\boldsymbol{Y}} \\ \boldsymbol{W}(\boldsymbol{\theta}) &= (\boldsymbol{G}_Y(\boldsymbol{\theta})\boldsymbol{G}_Y(\boldsymbol{\theta})^T)^{-1} \end{aligned}$$

Consider the differences between Eqs. (4.8)–(C.12), (4.19)–(C.15) and Eqs. (4.30)–(C.18). Notice how the correlation and covariance matrices do not have to be inverted. Instead the only operation that has to be computed on this matrices is the

Cholesky factorisation. Since now this new equations are considerably more stable, the formulation and the involved computational effort are considerably lessened.

C.1.5 Module 4

The last module of the MBA can also be transformed with the above mentioned factorisations. The equations for each of the components of Eq. (4.4) will be detailed below.

Physics-based model

With the above mentioned simplifications applied, the expected value and covariance matrix of Section 4.5.5 can be written as

$$\mathrm{E}[\boldsymbol{y}^m(\boldsymbol{x}_*)|\boldsymbol{\theta},\boldsymbol{D}^m,\hat{\boldsymbol{\phi}}^m] \;\; = \boldsymbol{h}_m(\boldsymbol{x}_*,\boldsymbol{\theta})\hat{\boldsymbol{\beta}}_m + \widetilde{\mathcal{R}}_m(\boldsymbol{x}_*,\boldsymbol{\theta})^T(\widetilde{\boldsymbol{Y}}^m - \widetilde{\boldsymbol{H}}_m\hat{\boldsymbol{\beta}}_m) \tag{C.19}$$

$$\begin{aligned}\boldsymbol{V}[\boldsymbol{y}^m(\boldsymbol{x}_*)^T,\boldsymbol{y}^m(\boldsymbol{x}'_*)^T|\boldsymbol{\theta},\boldsymbol{D}^m,\hat{\boldsymbol{\phi}}^m] = \hat{\boldsymbol{\Sigma}}_m^2 \otimes \{R_m((\boldsymbol{x}_*,\boldsymbol{\theta}),(\boldsymbol{x}'_*,\boldsymbol{\theta}))\\ -\widetilde{\mathcal{R}}_m(\boldsymbol{x}_*,\boldsymbol{\theta})^T\widetilde{\mathcal{R}}_m(\boldsymbol{x}'_*,\boldsymbol{\theta}) + \widetilde{\boldsymbol{u}}(\boldsymbol{x}_*,\boldsymbol{\theta})^T(\boldsymbol{G}_m\boldsymbol{G}_m^T)^{-1}\widetilde{\boldsymbol{u}}(\boldsymbol{x}'_*,\boldsymbol{\theta})\}\end{aligned} \tag{C.20}$$

where

$$\boldsymbol{C}_m\widetilde{\mathcal{R}}_m(\boldsymbol{x}_*,\boldsymbol{\theta}) = \mathcal{R}_m(\boldsymbol{x}_*,\boldsymbol{\theta}) \tag{C.21}$$

$$\widetilde{\boldsymbol{u}}(\boldsymbol{x}_*,\boldsymbol{\theta}) = \boldsymbol{h}_m(\boldsymbol{x}_*,\boldsymbol{\theta})^T - \widetilde{\boldsymbol{H}}_m^T\widetilde{\mathcal{R}}_m(\boldsymbol{x}_*,\boldsymbol{\theta}). \tag{C.22}$$

In compact form these equations can be written as

$$\mathrm{E}[\boldsymbol{y}^m(\boldsymbol{x}_*)|\boldsymbol{\theta},\boldsymbol{D}^m,\hat{\boldsymbol{\phi}}^m] \;\; = \boldsymbol{h}_m\hat{\boldsymbol{\beta}}_m + \widetilde{\mathcal{R}}_m^T(\widetilde{\boldsymbol{Y}}^m - \widetilde{\boldsymbol{H}}_m\hat{\boldsymbol{\beta}}_m) \tag{C.23}$$

$$\boldsymbol{V}[\boldsymbol{y}^m(\boldsymbol{x}_*)^T,\boldsymbol{y}^m(\boldsymbol{x}'_*)^T|\boldsymbol{\theta},\boldsymbol{D}^m,\hat{\boldsymbol{\phi}}^m] = \hat{\boldsymbol{\Sigma}}_m^2 \otimes \{R_m - \widetilde{\mathcal{R}}_m^T\widetilde{\mathcal{R}}_m + \widetilde{\boldsymbol{u}}^T(\boldsymbol{G}_m\boldsymbol{G}_m^T)^{-1}\widetilde{\boldsymbol{u}}\} \tag{C.24}$$

Calculating the unconditional posterior distribution requires marginalising the above moments with respect to the posterior distribution from Module 3.

$$\mathrm{E}[\boldsymbol{y}^m(\boldsymbol{x}_*)^T|\boldsymbol{D}^m,\hat{\boldsymbol{\phi}}^m] = \mathrm{E}[\mathrm{E}[\boldsymbol{y}^m(\boldsymbol{x}_*)^T|\boldsymbol{\theta},\boldsymbol{D}^m,\hat{\boldsymbol{\phi}}^m]] \tag{C.25}$$

$$
\begin{aligned}
V[\boldsymbol{y}^m(\boldsymbol{x}_*)^T\boldsymbol{y}^m(\boldsymbol{x}'_*)^T|\boldsymbol{D}^m,\hat{\boldsymbol{\phi}}^m] &= \mathrm{E}[V[\boldsymbol{y}^m(\boldsymbol{x}_*)^T\boldsymbol{y}^m(\boldsymbol{x}'_*)^T|\boldsymbol{\theta},\boldsymbol{D}^m,\hat{\boldsymbol{\phi}}^m]] \\
&+ V[\mathrm{E}[\boldsymbol{y}^m(\boldsymbol{x}_*)^T|\boldsymbol{\theta},\boldsymbol{D}^m,\hat{\boldsymbol{\phi}}^m],\mathrm{E}[\boldsymbol{y}^m(\boldsymbol{x}'_*)^T|\boldsymbol{\theta},\boldsymbol{D}^m,\hat{\boldsymbol{\phi}}^m]]
\end{aligned} \tag{C.26}
$$

Due to the complexity of the posterior distribution of $\boldsymbol{\theta}$ these integrals have to be computed with numerical methods. For the present implementation Gauss Legendre or Monte Carlo techniques have been used for one dimensional or multi dimensional examples, respectively.

Experimental response

With the above mentioned simplifications applied the expected value and covariance matrix of the experimental response of Section 4.5.5 can be written in compact form as

$$
\mathrm{E}[\boldsymbol{y}^e(\boldsymbol{x}_*)^T|\boldsymbol{\theta},\boldsymbol{D},\hat{\boldsymbol{\phi}}] = \boldsymbol{h}_Y\hat{\boldsymbol{\beta}} + \widetilde{\mathcal{V}}^T(\widetilde{\boldsymbol{Y}} - \widetilde{\boldsymbol{H}}_Y\hat{\boldsymbol{\beta}}) \tag{C.27}
$$

$$
\begin{aligned}
\boldsymbol{V}[\boldsymbol{y}^e(\boldsymbol{x}_*)^T,\boldsymbol{y}^e(\boldsymbol{x}'_*)^T|\boldsymbol{\theta},\boldsymbol{D},\hat{\boldsymbol{\phi}}] &= \hat{\boldsymbol{\Sigma}}_m^2 \otimes R_m + \hat{\boldsymbol{\Sigma}}_\delta^2 \otimes R_\delta + \hat{\boldsymbol{\Lambda}} \otimes \boldsymbol{I}_s \\
&- \widetilde{\mathcal{V}}^T\widetilde{\mathcal{V}} + \widetilde{\boldsymbol{U}}^T(\boldsymbol{G}_Y\boldsymbol{G}_Y^T)^{-1}\widetilde{\boldsymbol{U}}
\end{aligned} \tag{C.28}
$$

where the transformed global relational covariance matrix is defined as

$$
\boldsymbol{C}_Y\widetilde{\mathcal{V}}(\boldsymbol{x}_*,\boldsymbol{\theta}) = \mathcal{V}(\boldsymbol{x}_*,\boldsymbol{\theta}), \tag{C.29}
$$

and

$$
\boldsymbol{h}_Y(\boldsymbol{x}_*,\boldsymbol{\theta}) = [\boldsymbol{I}_q \otimes \boldsymbol{h}_m(\boldsymbol{x}_*,\boldsymbol{\theta}) \quad \boldsymbol{I}_q \otimes \boldsymbol{h}_\delta(\boldsymbol{x}_*)] \tag{C.30}
$$

$$
\widetilde{\boldsymbol{U}}(\boldsymbol{x}_*,\boldsymbol{\theta}) = \boldsymbol{h}_Y(\boldsymbol{x}_*,\boldsymbol{\theta})^T - \widetilde{\boldsymbol{H}}_Y(\boldsymbol{\theta})^T\widetilde{\mathcal{V}}(\boldsymbol{x}_*,\boldsymbol{\theta}) \tag{C.31}
$$

Similarly to the previous section the *unconditional* moments can be calculated from the law of total expectation and law of total variance.

Discrepancy function

For simplicity and computational reasons the discrepancy function shall be calculated simply as

$$
\mathrm{E}[\boldsymbol{\delta}(\boldsymbol{x}_*)|\boldsymbol{D},\hat{\boldsymbol{\phi}^\delta}] = \mathrm{E}[\boldsymbol{y}^e(\boldsymbol{x}_*)|\boldsymbol{D},\hat{\boldsymbol{\phi}}] - \mathrm{E}[\boldsymbol{y}^m(\boldsymbol{x}_*)|\boldsymbol{D}^m,\hat{\boldsymbol{\phi}^m}] \tag{C.32}
$$

since the dependency of $\boldsymbol{\theta}$ has already been explicited in the physycs-based model and experimental response of the two previous sections. This decision is justified simply to avoid an additional integration over the domain of the structural parameters $\boldsymbol{\theta}$. The conditional covariance matrix can be computed as

$$\begin{aligned}\boldsymbol{V}[\boldsymbol{\delta}(\boldsymbol{x}_*)\boldsymbol{\delta}(\boldsymbol{x}'_*)|\boldsymbol{\theta},\boldsymbol{D},\hat{\phi}] = \quad & \hat{\boldsymbol{\Sigma}}_\delta^2 \otimes \{R_\delta(\boldsymbol{x}_*,\boldsymbol{x}'_*) - \widetilde{\mathcal{R}}_\delta(\boldsymbol{x}_*)^T\widetilde{\mathcal{R}}_\delta(\boldsymbol{x}'_*) \\ & +(\boldsymbol{h}_\delta(\boldsymbol{x}_*)^T - \widetilde{\boldsymbol{H}}_\delta^T\widetilde{\mathcal{R}}_\delta(\boldsymbol{x}_*))^T(\boldsymbol{G}_\delta\boldsymbol{G}_\delta^T)^{-1}(\boldsymbol{h}_\delta(\boldsymbol{x}'_*)^T - \widetilde{\boldsymbol{H}}_\delta^T\widetilde{\mathcal{R}}_\delta(\boldsymbol{x}'_*))\},\end{aligned} \tag{C.33}$$

where

$$\boldsymbol{C}_\delta\widetilde{\mathcal{R}}_\delta(\boldsymbol{x}_*) = \mathcal{R}_\delta(\boldsymbol{x}_*), \tag{C.34}$$

and similarly to the previous section the *unconditional* moments can be calculated from the law of total expectation and law of total variance.

C.2 Data transformation into standardised units

It is often very useful to standardise the datasets to improve the numerical performance, avoid rounding errors and other numerical issues. The user of the code does not have to be concerned with this operation, since it is done automatically as a first step, before processing the datasets as

$$\boldsymbol{X}_z^m = \frac{\boldsymbol{X}^m - \boldsymbol{m}_X}{\boldsymbol{\sigma}_X} \qquad \boldsymbol{Y}_z^m = \frac{\boldsymbol{Y}^m - \boldsymbol{m}_Y}{\boldsymbol{\sigma}_Y}, \tag{C.35}$$

where $\boldsymbol{m}_X$, $\boldsymbol{m}_Y$, $\boldsymbol{\sigma}_X$ and $\boldsymbol{\sigma}_Y$ represent means of the design variables/structural parameters, measured responses, standard deviations of the design variables/structural parameters and measured responses, respectively.

The only two remarks are

- that both the datasets, from the physics-based model $\boldsymbol{D}_m$ and observations $\boldsymbol{D}_e$ have to be standardised with the same conditions. In this case the physics-based model was used as the reference;
- since the discrepancy function represents a difference between two components i.e. $\boldsymbol{\delta} = \boldsymbol{Y}^e - \hat{\boldsymbol{Y}}^m$ to transform it back into the original scale it suffices to multiply it by the variance scaling factors, i.e. $\boldsymbol{\sigma}_X, \boldsymbol{\sigma}_Y$

C.3 Positive-definitess of covariance matrices

It is important to ensure that the covariance matrices and correlation matrices used throughout the formulation are always positive-definite. The Cholesky transforma-

tions shown above should guarantee this requirement in most of situations. However it should be noted that an inappropriate choice of the correlation functions of the physics-based model and the discrepancy function can result in non-positive definite matrices, particularly during the numerical integration of module 2. For this reason, it is important to check the condition numbers of these matrices, and whether they are positive definite. If not, it is advisable to try a different correlation function from the available options in Table 3.2.

Bibliography

[1] Ko, J., AND Ni, Y. Technology Developments in Structural Health Monitoring of Large-Scale Bridges. *Engineering Structures* 27, 12 (Oct. 2005), 1715–1725. ISSN: 01410296. DOI: 10.1016/j.engstruct.2005.02.021.

[2] *Structural Health Monitoring Market: Global Industry Analysis and Opportunity Assessment, 2016-2026*. 4394409. Future Market Insight Global & Consulting Pvt Ltd, Oct. 2016, 207.

[3] Inaudi, D. 20 Years of SHM. In: *8th European Workshop on Structural Health Monitoring, EWSHM 2016*. Vol. 4. 2016, 2534–2542.

[4] Figueiredo, E., Moldovan, I., AND J.M. Barata Marques, M. *Condition Assessment of Bridges: Past, Present and Future. A Complementary Approach.* Dec. 2013. ISBN: 978-972-54-0402-7.

[5] Farrar, C. R., Doebling, S. W., AND Nix, D. A. Vibration-Based Structural Damage Identification. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 359, 1778 (15th Jan. 2001), 131–149. ISSN: 1364-503X, 1471-2962. DOI: 10.1098/rsta.2000.0717.

[6] Brownjohn, J. M. W., AND Xia, P.-Q. Dynamic Assessment of Curved Cable-Stayed Bridge by Model Updating. *Journal of Structural Engineering* 126, 2 (Feb. 2000), 252–260. ISSN: 0733-9445, 1943-541X. DOI: 10.1061/(ASCE)0733-9445(2000)126:2(252).

[7] Farrar, C. R., AND Worden, K. An Introduction to Structural Health Monitoring. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 365, 1851 (15th Feb. 2007), 303–315. ISSN: 1364-503X, 1471-2962. DOI: 10.1098/rsta.2006.1928.

[8] Barthorpe, R. On Model- and Data-Based Approaches to Structural Health Monitoring. PhD. Sheffield, 28th Jan. 2011. 237 pp.

[9] Rytter, A. Vibrational Based Inspection of Civil Engineering Structures. Denmark: Dept. of Building Technology and Structural Engineering, Aalborg University, 1993.

[10] Farrar, C. R., AND Worden, K. *Structural Health Monitoring: A Machine Learning Perspective*. John Wiley & Sons, Ltd, Chichester, UK, 1st Nov. 2012. ISBN: 978-1-118-44311-8 978-1-119-99433-6. DOI: 10.1002/9781118443118.

[11] Sirca, G., AND Adeli, H. System Identification in Structural Engineering. *Scientia Iranica* 19, 6 (Dec. 2012), 1355–1364. ISSN: 10263098. DOI: 10.1016/j.scient.2012.09.002.

[12] Catbas, F. N., Kijewski-Correa, T. L., Aktan, A. E., AND Structural Engineering Institute, eds. *Structural Identification of Constructed Systems: Approaches, Methods, and Technologies for Effective Practice of St-Id*. American Society of Civil Engineers, Reston, Virginia, 2013. 234 pp. ISBN: 978-0-7844-1197-1.

[13] Guratzsch, R. F. Sensor Placement Optimization under Uncertainty for Structural Health Monitoring Systems of Hot Aerospace Structures. PhD. Nashville, Tennessee: Vanderbilt University, 2007.

[14] Spiridonakos, M., Chatzi, E., AND Sudret, B. Polynomial Chaos Expansion Models for the Monitoring of Structures under Operational Variability. *ASCE ASME Journal of Risk and Uncertainty in Engineering Systems* 2, 3 (Sept. 2016), B4016003–.

[15] Laory, I., Hadj Ali, N. B., Trinh, T. N., AND Smith, I. F. C. Measurement System Configuration for Damage Identification of Continuously Monitored Structures. *Journal of Bridge Engineering* 17, 6 (Nov. 2012), 857–866. ISSN: 1084-0702, 1943-5592. DOI: 10.1061/(ASCE)BE.1943-5592.0000386.

[16] Vestroni, F. Structural Identification and Damage Detection. In: *Dynamic Methods for Damage Detection in Structures*. Ed. by Morassi, A., AND Vestroni, F. Springer Vienna, Vienna, 2008, 111–135. ISBN: 978-3-211-78777-9. DOI: 10.1007/978-3-211-78777-9_5.

[17] Melchers, R. E. *Structural Reliability Analysis and Prediction*. 2nd ed. John Wiley, Chichester ; New York, 1999. 437 pp. ISBN: 978-0-471-98771-0 978-0-471-98324-8.

[18] Jaynes, E. T., AND Bretthorst, G. L. *Probability Theory: The Logic of Science*. Cambridge University Press, Cambridge, UK ; New York, NY, 2003. 727 pp. ISBN: 978-0-521-59271-0.

[19] Box, G. E. P. Science and Statistics. *Journal of the American Statistical Association* 71, 356 (Dec. 1976), 791–799. ISSN: 0162-1459, 1537-274X. DOI: 10.1080/01621459.1976.10480949.

[20] Beck, J., AND Katafygiotis, L. Updating Models and Their Uncertainties. I: Bayesian Statistical Framework. *Journal of Engineering Mechanics* 124, 4 (1998), 455–461. DOI: 10.1061/(ASCE)0733-9399(1998)124:4(455).

[21] Legault, J., Langley, R., AND Woodhouse, J. Physical Consequences of a Nonparametric Uncertainty Model in Structural Dynamics. *Journal of Sound and Vibration* 331, 25 (Dec. 2012), 5469–5487. ISSN: 0022460X. DOI: 10.1016/j.jsv.2012.07.017.

[22] Helton, J. C. Treatment of Uncertainty in Performance Assessments for Complex Systems. *Risk Analysis* 14, 4 (Aug. 1994), 483–511. ISSN: 0272-4332, 1539-6924. DOI: 10.1111/j.1539-6924.1994.tb00266.x.

[23] Helton, J. Uncertainty and Sensitivity Analysis in the Presence of Stochastic and Subjective Uncertainty. *Journal of Statistical Computation and Simulation* 57, 1-4 (Apr. 1997), 3–76. ISSN: 0094-9655, 1563-5163. DOI: 10.1080/00949659708811803.

[24] Hoffman, F. O., AND Hammonds, J. S. Propagation of Uncertainty in Risk Assessments: The Need to Distinguish Between Uncertainty Due to Lack of Knowledge and Uncertainty Due to Variability. *Risk Analysis* 14, 5 (Oct. 1994), 707–712. ISSN: 0272-4332, 1539-6924. DOI: 10.1111/j.1539-6924.1994.tb00281.x.

[25] Parry, G. W. The Characterization of Uncertainty in Probabilistic Risk Assessments of Complex Systems. *Reliability Engineering & System Safety* 54, 2-3 (Nov. 1996), 119–126. ISSN: 09518320. DOI: 10.1016/S0951-8320(96)00069-5.

[26] *Sensitivity Analysis Wikipedia*. 25th May 2015. URL: https://en.wikipedia.org/wiki/Sensitivity_analysis.

[27] Wan, H.-P., AND Ren, W.-X. Parameter Selection in Finite-Element-Model Updating by Global Sensitivity Analysis Using Gaussian Process Metamodel. *JOURNAL OF STRUCTURAL ENGINEERING* 141, 6 (June 2015). ISSN: 0733-9445. DOI: 10.1061/(ASCE)ST.1943-541X.0001108.

[28] Wan, H.-P., AND Ren, W.-X. A Residual-Based Gaussian Process Model Framework for Finite Element Model Updating. *Computers & Structures* 156 (Aug. 2015), 149–159. ISSN: 00457949. DOI: 10.1016/j.compstruc.2015.05.003.

[29] Wan, H.-P., Mao, Z., Todd, M. D., AND Ren, W.-X. Analytical Uncertainty Quantification for Modal Frequencies with Structural Parameter Uncertainty Using a Gaussian Process Metamodel. *Engineering Structures* 75 (Sept. 2014), 577–589. ISSN: 01410296. DOI: 10.1016/j.engstruct.2014.06.028.

[30] García-Macías, E., Castro-Triguero, R., Friswell, M. I., Adhikari, S., AND Sáez, A. Metamodel-Based Approach for Stochastic Free Vibration Analysis of Functionally Graded Carbon Nanotube Reinforced Plates. *Composite Structures* 152 (Sept. 2016), 183–198. ISSN: 02638223. DOI: 10.1016/j.compstruct.2016.05.019.

[31] Ghasemi, H., Rafiee, R., Zhuang, X., Muthu, J., AND Rabczuk, T. Uncertainties Propagation in Metamodel-Based Probabilistic Optimization of CNT/Polymer Composite Structure Using Stochastic Multi-Scale Modeling. *Computational Materials Science* 85 (Apr. 2014), 295–305. ISSN: 09270256. DOI: 10.1016/j.commatsci.2014.01.020.

[32] Haldar, A., AND Mahadevan, S. *Probability, Reliability, and Statistical Methods in Engineering Design*. John Wiley, New York ; Chichester [England], 2000. 304 pp. ISBN: 978-0-471-33119-3.

[33] Ghanem, R., AND Spanos, P. D. Polynomial Chaos in Stochastic Finite Elements. *Journal of Applied Mechanics* 57, 1 (1990), 197. ISSN: 00218936. DOI: 10.1115/1.2888303.

[34] Buhmann, M. D. *Radial Basis Functions: Theory and Implementations*. Transferred to digital print. Cambridge monographs on applied and computational mathematics 12. Cambridge Univ. Press, Cambridge, 2006. 259 pp. ISBN: 978-0-521-10133-2 978-0-521-63338-3.

[35] Wiberg, J., Karoumi, R., AND Pacoste, C. Statistical Screening of Individual and Joint Effect of Several Modelling Factors on the Dynamic Finite Element Response of a Railway Bridge. *Computers & Structures* 106-107 (Sept. 2012), 91–104. ISSN: 00457949. DOI: 10.1016/j.compstruc.2012.04.008.

[36] Jesus, A., Dimitrovová, Z., AND Silva, M. The Qualitative Influence of Model Parameters on the Dynamic Response of a Viaduct. *International Journal of*

Railway Technology 1, 2 (2012), 89–113. ISSN: 2049-5358. DOI: 10.4203/ijrt.1.2.5.

[37] Jesus, A. H., Dimitrovová, Z., AND Silva, M. A. A Statistical Analysis of the Dynamic Response of a Railway Viaduct. *Engineering Structures* 71 (July 2014), 244–259. ISSN: 01410296. DOI: 10.1016/j.engstruct.2014.04.012.

[38] Box, G. E. P., Hunter, J. S., AND Hunter, W. G. *Statistics for Experimenters: Design, Innovation, and Discovery*. 2nd ed. Wiley series in probability and statistics. Wiley-Interscience, Hoboken, N.J, 2005. 633 pp. ISBN: 978-0-471-71813-0.

[39] Arendt, P. D., Apley, D. W., AND Chen, W. Quantification of Model Uncertainty: Calibration, Model Discrepancy, and Identifiability. *Journal of Mechanical Design* 134, 10 (Oct. 2012). ISSN: 1050-0472. DOI: 10.1115/1.4007390.

[40] De Laplace, P. *Théorie Analytique Des Probabilités*. Ve. Courcier, 1812.

[41] Bayes, M., AND Price, M. An Essay towards Solving a Problem in the Doctrine of Chances. By the Late Rev. Mr. Bayes, F. R. S. Communicated by Mr. Price, in a Letter to John Canton, A. M. F. R. S. *Philosophical Transactions of the Royal Society of London* 53, 0 (1st Jan. 1763), 370–418. ISSN: 0261-0523. DOI: 10.1098/rstl.1763.0053.

[42] Zapata-Vázquez, R. E., O'Hagan, A., AND Soares Bastos, L. Eliciting Expert Judgements about a Set of Proportions. *Journal of Applied Statistics* 41, 9 (2nd Sept. 2014), 1919–1933. ISSN: 0266-4763, 1360-0532. DOI: 10.1080/02664763.2014.898131.

[43] Greenland, S. Bayesian Perspectives for Epidemiological Research: I. Foundations and Basic Methods. *International Journal of Epidemiology* 35, 3 (1st June 2006), 765–775. ISSN: 1464-3685, 0300-5771. DOI: 10.1093/ije/dyi312.

[44] Tikhonov, A. N., AND Arsenin, V. Y. *Solutions of Ill-Posed Problems*. Scripta series in mathematics. Winston ; distributed solely by Halsted Press, Washington : New York, 1977. 258 pp. ISBN: 978-0-470-99124-4.

[45] Casella, G., AND Berger, R. L. *Statistical Inference*. 2nd ed. Thomson Learning, Australia ; Pacific Grove, CA, 2002. 660 pp. ISBN: 978-0-534-24312-8.

[46] Sankararaman, S., AND Mahadevan, S. Integration of Model Verification, Validation, and Calibration for Uncertainty Quantification in Engineering Systems. *Reliability Engineering & System Safety* 138, 0 (2015), 194 –209. ISSN: 0951-8320. DOI: 10.1016/j.ress.2015.01.023.

[47] Ljung, L., AND Glad, T. On Global Identifiability for Arbitrary Model Parametrizations. *Automatica* 30, 2 (Feb. 1994), 265–276. ISSN: 00051098. DOI: 10.1016/0005-1098(94)90029-9.

[48] Katafygiotis, L. S., AND Beck, J. L. Updating Models and Their Uncertainties. II: Model Identifiability. *Journal of Engineering Mechanics* 124, 4 (Apr. 1998), 463–467. ISSN: 0733-9399, 1943-7889. DOI: 10.1061/(ASCE)0733-9399(1998)124:4(463).

[49] Yuen, K.-V. Bayesian Methods for Structural Dynamics and Civil Engineering. In: *Bayesian Methods for Structural Dynamics and Civil Engineering*. John Wiley & Sons, Ltd, 2010, 1–10. ISBN: 978-0-470-82456-6.

[50] Kay, S. M. *Fundamentals of Statistical Signal Processing*. Prentice Hall signal processing series. Prentice-Hall PTR, Englewood Cliffs, N.J, 1993. 3 pp. ISBN: 978-0-13-345711-7 978-0-13-504135-2 978-0-13-280803-3.

[51] Sohn, H., AND Law, K. H. A Bayesian Probabilistic Approach for Structure Damage Detection. *Dyn* 26 (1997), 1259–1281.

[52] Papadimitriou, C., Beck, J., AND Katafygiotis, L. Updating Robust Reliability Using Structural Test Data. *Probabilistic Engineering Mechanics* 16, 2 (Apr. 2001), 103–113. ISSN: 02668920. DOI: 10.1016/S0266-8920(00)00012-6.

[53] Beck, J. Robust Stochastic Predictions of Dynamic Response During and Monitoring of Structures. *Procedia Engineering* 14 (2011), 32–40. ISSN: 18777058. DOI: 10.1016/j.proeng.2011.07.004.

[54] Simoen, E., Roeck, G. D., AND Lombaert, G. Dealing with Uncertainty in Model Updating for Damage Assessment: A Review. *Mechanical Systems and Signal Processing* 56–57, 0 (2015), 123 –149. ISSN: 0888-3270. DOI: 10.1016/j.ymssp.2014.11.001.

[55] Zuev, K. M., Beck, J. L., Au, S.-K., AND Katafygiotis, L. S. Bayesian Post-Processor and Other Enhancements of Subset Simulation for Estimating Failure Probabilities in High Dimensions. *COMPUTERS & STRUCTURES* 92-93 (Feb. 2012), 283–296. ISSN: 0045-7949. DOI: 10.1016/j.compstruc.2011.10.017.

[56] Beck, J., AND Au, S.-K. Bayesian Updating of Structural Models and Reliability Using Markov Chain Monte Carlo Simulation. *Journal of Engineering Mechanics* 128, 4 (2002), 380–391. DOI: 10.1061/(ASCE)0733-9399(2002)128:4(380).

[57] Cheung, S., AND Beck, J. Bayesian Model Updating Using Hybrid Monte Carlo Simulation with Application to Structural Dynamic Models with Many Uncertain Parameters. *Journal of Engineering Mechanics* 135, 4 (2009), 243–255. DOI: 10.1061/(ASCE)0733-9399(2009)135:4(243).

[58] Saito, T., AND Beck, J. L. Bayesian Model Selection for ARX Models and Its Application to Structural Health Monitoring. *Earthquake Engineering & Structural Dynamics* 39, 15 (Dec. 2010), 1737–1759. ISSN: 00988847. DOI: 10.1002/eqe.1006.

[59] Yuen, K., AND Beck, J. Updating Properties of Nonlinear Dynamical Systems with Uncertain Input. *Journal of Engineering Mechanics* 129, 1 (2003), 9–20. DOI: 10.1061/(ASCE)0733-9399(2003)129:1(9).

[60] Kuok, S.-C., AND Yuen, K.-V. Structural Health Monitoring of Canton Tower Using Bayesian Framework. *SMART STRUCTURES AND SYSTEMS* 10, 4-5 (OCT-NOV 2012), 375–391. ISSN: 1738-1584.

[61] Yuen, K.-V., Beck, J. L., AND Katafygiotis, L. S. Efficient Model Updating and Health Monitoring Methodology Using Incomplete Modal Data without Mode Matching. *Structural Control and Health Monitoring* 13, 1 (2006), 91–107. ISSN: 1545-2263. DOI: 10.1002/stc.144.

[62] Beck, J., AND Yuen, K. Model Selection Using Response Measurements: Bayesian Probabilistic Approach. *Journal of Engineering Mechanics* 130, 2 (2004), 192–203. DOI: 10.1061/(ASCE)0733-9399(2004)130:2(192).

[63] Yuen, K.-V., Katafygiotis, L. S., Papadimitriou, C., AND Mickleborough, N. C. Optimal Sensor Placement Methodology for Identification with Unmeasured Excitation. *Journal of Dynamic Systems, Measurement, and Control* 123, 4 (2001), 677. ISSN: 00220434. DOI: 10.1115/1.1410929.

[64] Katafygiotis, L. S., AND Yuen, K.-V. Bayesian Spectral Density Approach for Modal Updating Using Ambient Data. *Earthquake Engineering & Structural Dynamics* 30, 8 (Aug. 2001), 1103–1123. ISSN: 0098-8847, 1096-9845. DOI: 10.1002/eqe.53.

[65] Yuen, K.-V., AND Katafygiotis, L. Bayesian Fast Fourier Transform Approach for Modal Updating Using Ambient Data. *Advances in Structural Engineering* 6, 2 (May 2003), 81–95. ISSN: 1369-4332. DOI: 10.1260/136943303769013183.

[66] Yuen, K.-V., AND Katafygiotis, L. S. Bayesian Time–domain Approach for Modal Updating Using Ambient Data. *Probabilistic Engineering Mechanics* 16, 3 (July 2001), 219–231. ISSN: 02668920. DOI: 10.1016/S0266-8920(01)00004-2.

[67] Beck, J. L., Chan, E., AND Papadimitriou, C. Statistical Methodology for Optimal Sensor Locations for Damage Detection in Structures. In: *Proceedings 16th International Modal Analysis Conference*. 1998.

[68] Katafygiotis, L. S., Papadimitriou, C., AND Lam, H.-F. A Probabilistic Approach to Structural Model Updating. *Soil Dynamics and Earthquake Engineering* 17, 7-8 (Oct. 1998), 495–507. ISSN: 02677261. DOI: 10.1016/S0267-7261(98)00008-6.

[69] Papadimitriou, C., Beck, J. L., AND Au, S.-K. Entropy-Based Optimal Sensor Location for Structural Model Updating. *Journal of Vibration and Control* 6, 5 (2000), 781–800. DOI: 10.1177/107754630000600508.

[70] Lehmann, E. L., AND Casella, G. *Theory of Point Estimation*. 2nd ed. Springer texts in statistics. Springer, New York, 1998. 589 pp. ISBN: 978-0-387-98502-2.

[71] Chow, H. M., Lam, H. F., Yin, T., AND Au, S. K. Optimal Sensor Configuration of a Typical Transmission Tower for the Purpose of Structural Model Updating. *Structural Control & Health Monitoring* 18, 3 (Apr. 2011), 305–320. ISSN: 1545-2255. DOI: 10.1002/stc.372.

[72] Simoen, E., Papadimitriou, C., AND Lombaert, G. On Prediction Error Correlation in Bayesian Model Updating. *Journal of Sound and Vibration* 332, 18 (Sept. 2013), 4136–4152. ISSN: 0022460X. DOI: 10.1016/j.jsv.2013.03.019.

[73] Spiridonakos, M., AND Chatzi, E. Metamodeling of Dynamic Nonlinear Structural Systems through Polynomial Chaos NARX Models. *Computers & Structures* 157 (Sept. 2015), 99–113. ISSN: 0045-7949. DOI: 10.1016/j.compstruc.2015.05.002.

[74] Kang, F., Li, J.-j., AND Xu, Q. Virus Coevolution Partheno-Genetic Algorithms for Optimal Sensor Placement. *Advanced Engineering Informatics* 22, 3 (July 2008), 362–370. ISSN: 14740346. DOI: 10.1016/j.aei.2008.02.001.

[75] Rao, A. R. M., AND Anandakumar, G. Optimal Placement of Sensors for Structural System Identification and Health Monitoring Using a Hybrid Swarm Intelligence Technique. *Smart Materials and Structures* 16, 6 (2007), 2658.

[76] Sun, H., AND Büyüköztürk, O. Optimal Sensor Placement in Structural Health Monitoring Using Discrete Optimization. *Smart Materials and Structures* 24, 12 (2015), 125034.

[77] Goldberg, D. E. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley Pub. Co, Reading, Mass, 1989. 412 pp. ISBN: 978-0-201-15767-3.

[78] Chan, E. Optimal Design of Building Structures Using Genetic Algorithms (Jan. 1997).

[79] Au, S.-K. *Operational Modal Analysis*. Springer Berlin Heidelberg, New York, NY, 2017. ISBN: 978-981-10-4117-4.

[80] Au, S.-K. Fast Bayesian FFT Method for Ambient Modal Identification with Separated Modes. *Journal of Engineering Mechanics* 137, 3 (Mar. 2011), 214–226. ISSN: 0733-9399, 1943-7889. DOI: 10.1061/(ASCE)EM.1943-7889.0000213.

[81] Au, S.-K. Erratum for "Fast Bayesian FFT Method for Ambient Modal Identification with Separated Modes". *Journal of Engineering Mechanics* 137, 3 (Mar. 2011), 214–226. ISSN: 0733-9399, 1943-7889. DOI: 10.1061/(ASCE)EM.1943-7889.0000213.

[82] Au, S.-K., Zhang, F.-L., AND Ni, Y.-C. Bayesian Operational Modal Analysis: Theory, Computation, Practice. *Computers & Structures* 126, 0 (2013), 3 – 14. ISSN: 0045-7949. DOI: 10.1016/j.compstruc.2012.12.015.

[83] Au, S.-K. Uncertainty Law in Ambient Modal Identification—Part I: Theory. *Mechanical Systems and Signal Processing* 48, 1–2 (2014), 15 –33. ISSN: 0888-3270. DOI: 10.1016/j.ymssp.2013.07.016.

[84] Au, S.-K. Uncertainty Law in Ambient Modal Identification—Part II: Implication and Field Verification. *Mechanical Systems and Signal Processing* 48, 1–2 (2014), 34 –48. ISSN: 0888-3270. DOI: 10.1016/j.ymssp.2013.07.017.

[85] Au, S.-K., AND Ni, Y.-C. Fast Bayesian Modal Identification of Structures Using Known Single-Input Forced Vibration Data. *Structural Control and Health Monitoring* 21, 3 (2014), 381–402. ISSN: 1545-2263. DOI: 10.1002/stc.1571.

[86] Ntotsios, E., Papadimitriou, C., Panetsos, P., Karaiskos, G., Perros, K., AND Perdikaris, P. C. Bridge Health Monitoring System Based on Vibration Measurements. *Bulletin of Earthquake Engineering* 7, 2 (May 2009), 469–483. ISSN: 1570-761X, 1573-1456. DOI: 10.1007/s10518-008-9067-4.

[87] Simoen, E., Moaveni, B., Conte, J., AND Lombaert, G. Uncertainty Quantification in the Assessment of Progressive Damage in a 7-Story Full-Scale Building Slice. *Journal of Engineering Mechanics* 139, 12 (2013), 1818–1830. DOI: 10.1061/(ASCE)EM.1943-7889.0000610.

[88] Ben Abdessalem, A., Dervilis, N., Wagg, D., AND Worden, K. Model Selection and Parameter Estimation in Structural Dynamics Using Approximate Bayesian Computation. *Mechanical Systems and Signal Processing* 99 (Jan. 2018), 306–325. ISSN: 08883270. DOI: 10.1016/j.ymssp.2017.06.017.

[89] Moaveni, B., He, X., Conte, J. P., AND Restrepo, J. I. Damage Identification Study of a Seven-Story Full-Scale Building Slice Tested on the UCSD-NEES Shake Table. *Structural Safety* 32, 5 (2010), 347 –356. ISSN: 0167-4730. DOI: http://dx.doi.org/10.1016/j.strusafe.2010.03.006.

[90] Moaveni, B., Conte, J. P., AND Hemez, F. M. Uncertainty and Sensitivity Analysis of Damage Identification Results Obtained Using Finite Element Model Updating. *Computer-Aided Civil and Infrastructure Engineering* 24, 5 (July 2009), 320–334. ISSN: 10939687, 14678667. DOI: 10.1111/j.1467-8667.2008.00589.x.

[91] Papadimitriou, C., AND Lombaert, G. The Effect of Prediction Error Correlation on Optimal Sensor Placement in Structural Dynamics. *Mechanical Systems and Signal Processing* 28 (Apr. 2012), 105–127. ISSN: 08883270. DOI: 10.1016/j.ymssp.2011.05.019.

[92] Simoen, E., Papadimitriou, C., De Roeck, G., AND Lombaert, G. Influence of the Prediction Error Correlation Model on Bayesian FE Model Updating Results. In: *Life-Cycle and Sustainability of Civil Infrastructure Systems - Proceedings of the 3rd International Symposium on Life-Cycle Civil Engineering, IALCCE 2012*. 2012, 192–199.

[93] Vincenzi, L., AND Simonini, L. Influence of Model Errors in Optimal Sensor Placement. *Journal of Sound and Vibration* 389 (Feb. 2017), 119–133. ISSN: 0022460X. DOI: 10.1016/j.jsv.2016.10.033.

[94] Vanik, M. W., Beck, J. L., AND Au, S. K. Bayesian Probabilistic Approach to Structural Health Monitoring. *Journal of Engineering Mechanics* 126, 7 (July 2000), 738–745. ISSN: 0733-9399, 1943-7889. DOI: 10.1061/(ASCE)0733-9399(2000)126:7(738).

[95] Matos, J. C., Cruz, P. J., Valente, I. B., Neves, L. C., AND Moreira, V. N. An Innovative Framework for Probabilistic-Based Structural Assessment with an Application to Existing Reinforced Concrete Structures. *Engineering Structures* 111 (Mar. 2016), 552–564. ISSN: 01410296. DOI: 10.1016/j.engstruct.2015.12.040.

[96] McFarland, J., AND Mahadevan, S. Multivariate Significance Testing and Model Calibration under Uncertainty. *Computer Methods in Applied Mechanics and Engineering* 197, 29-32 (May 2008), 2467–2479. ISSN: 00457825. DOI: 10.1016/j.cma.2007.05.030.

[97] Nishio, M., Marin, J., AND Fujino, Y. Uncertainty Quantification of the Finite Element Model of Existing Bridges for Dynamic Analysis. *Journal of Civil Structural Health Monitoring* 2, 3-4 (2012), 163–173. ISSN: 2190-5452. DOI: 10.1007/s13349-012-0026-z.

[98] Ebrahimian, H., Astroza, R., Conte, J. P., AND de Callafon, R. A. Nonlinear Finite Element Model Updating for Damage Identification of Civil Structures Using Batch Bayesian Estimation. *Mechanical Systems and Signal Processing* 84 (Feb. 2017), 194–222. ISSN: 08883270. DOI: 10.1016/j.ymssp.2016.02.002.

[99] Lam, H.-F., Yang, J., AND Au, S.-K. Bayesian Model Updating of a Coupled-Slab System Using Field Test Data Utilizing an Enhanced Markov Chain Monte Carlo Simulation Algorithm. *Engineering Structures* 102 (Nov. 2015), 144–155. ISSN: 01410296. DOI: 10.1016/j.engstruct.2015.08.005.

[100] Yin, T., Lam, H.-F., AND Chow, H.-M. A Bayesian Probabilistic Approach for Crack Characterization in Plate Structures: Crack Characterization in Plate Structures. *Computer-Aided Civil and Infrastructure Engineering* 25, 5 (10th Feb. 2010), 375–386. ISSN: 10939687, 14678667. DOI: 10.1111/j.1467-8667.2009.00647.x.

[101] Goller, B., AND Schueller, G. I. Investigation of Model Uncertainties in Bayesian Structural Model Updating. *Journal of Sound and Vibration* 330, 25 (5th Dec. 2011), 6122–6136. ISSN: 0022-460X. DOI: 10.1016/j.jsv.2011.07.036.

[102] Zheng, W., AND Yu, W. Probabilistic Approach to Assessing Scoured Bridge Performance and Associated Uncertainties Based on Vibration Measurements. *Journal of Bridge Engineering* 20, 6 (June 2015), 04014089. ISSN: 1084-0702, 1943-5592. DOI: 10.1061/(ASCE)BE.1943-5592.0000683.

[103] Ching, J., AND Beck, J. L. Bayesian Analysis of the Phase II IASC–ASCE Structural Health Monitoring Experimental Benchmark Data. *Journal of Engineering Mechanics* 130, 10 (Oct. 2004), 1233–1244. ISSN: 0733-9399, 1943-7889. DOI: 10.1061/(ASCE)0733-9399(2004)130:10(1233).

[104] Muto, M., AND Beck, J. L. Bayesian Updating and Model Class Selection for Hysteretic Structural Models Using Stochastic Simulation. *Journal of Vibration and Control* 14, 1-2 (Jan. 2008), 7–34. ISSN: 1077-5463, 1741-2986. DOI: 10.1177/1077546307079400.

[105] Hadjidoukas, P., Angelikopoulos, P., Papadimitriou, C., AND Koumoutsakos, P. Pi4U: A High Performance Computing Framework for Bayesian Uncertainty Quantification of Complex Models. *Journal of Computational Physics* 284 (Mar. 2015), 1–21. ISSN: 00219991. DOI: 10.1016/j.jcp.2014.12.006.

[106] Nagel, J. B., AND Sudret, B. A Unified Framework for Multilevel Uncertainty Quantification in Bayesian Inverse Problems. *Probabilistic Engineering Mechanics* 43 (Jan. 2016), 68–84. ISSN: 02668920. DOI: 10.1016/j.probengmech.2015.09.007.

[107] Christodoulou, K., AND Papadimitriou, C. Structural Identification Based on Optimally Weighted Modal Residuals. *Mechanical Systems and Signal Processing* 21, 1 (Jan. 2007), 4–23. ISSN: 08883270. DOI: 10.1016/j.ymssp.2006.05.011.

[108] Minson, S. E., Simons, M., Beck, J. L., Ortega, F., Jiang, J., Owen, S. E., Moore, A. W., Inbal, A., AND Sladen, A. Bayesian Inversion for Finite Fault Earthquake Source Models – II: The 2011 Great Tohoku-Oki, Japan Earthquake. *Geophysical Journal International* 198, 2 (2014), 922–940. DOI: 10.1093/gji/ggu170.

[109] Sankararaman, S., AND Mahadevan, S. Bayesian Methodology for Diagnosis Uncertainty Quantification and Health Monitoring. *Structural Control and Health Monitoring* 20, 1 (2013), 88–106. ISSN: 1545-2263. DOI: 10.1002/stc.476.

[110] Stull, C. J., Earls, C. J., AND Koutsourelakis, P.-S. Model-Based Structural Health Monitoring of Naval Ship Hulls. *Computer Methods in Applied Mechanics and Engineering* 200, 9-12 (Feb. 2011), 1137–1149. ISSN: 00457825. DOI: 10.1016/j.cma.2010.11.018.

[111] Erickson, G. J., Rychert, J. T., AND Smith, C. R., eds. *Maximum Entropy and Bayesian Methods: Boise, Idaho, U.S.A., 1997: Proceedings of the 17th International Workshop on Maximum Entropy and Bayesian Methods of Statistical Analysis*. Fundamental theories of physics v. 98. Kluwer Academic Publishers, Dordrecht ; Boston, 1998. 297 pp. ISBN: 978-0-7923-5047-7.

[112] Ferson, S., AND Ginzburg, L. R. Different Methods Are Needed to Propagate Ignorance and Variability. *Reliability Engineering & System Safety* 54, 2-3 (Nov. 1996), 133–144. ISSN: 09518320. DOI: 10.1016/S0951-8320(96)00071-3.

[113] Ferson, S., AND Oberkampf, W. L. Validation of Imprecise Probability Models. *International Journal of Reliability and Safety* 3 (1/2/3 2009), 3. ISSN: 1479-389X, 1479-3903. DOI: 10.1504/IJRS.2009.026832.

[114] Sohn, H. Effects of Environmental and Operational Variability on Structural Health Monitoring. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 365, 1851 (15th Feb. 2007), 539–560. ISSN: 1364-503X, 1471-2962. DOI: 10.1098/rsta.2006.1935.

[115] Doebling, S., Farrar, C., Prime, M., AND Shevitz, D. *Damage Identification and Health Monitoring of Structural and Mechanical Systems from Changes in Their Vibration Characteristics: A Literature Review*. LA–13070-MS, 249299. 1st May 1996. DOI: 10.2172/249299.

[116] Westgate, R. Environmental Effects on a Suspension Bridge's Performance. PhD. Sheffield, 2012.

[117] Shahabpoor, E., Pavic, A., Racic, V., AND Zivanovic, S. Effect of Group Walking Traffic on Dynamic Properties of Pedestrian Structures. *Journal of Sound and Vibration* 387 (Jan. 2017), 207–225. ISSN: 0022460X. DOI: 10.1016/j.jsv.2016.10.017.

[118] Pombo, J., Ambrósio, J., AND Silva, M. A New Wheel–rail Contact Model for Railway Dynamics. *Vehicle System Dynamics* 45, 2 (Feb. 2007), 165–189. ISSN: 0042-3114, 1744-5159. DOI: 10.1080/00423110600996017.

[119] Thor Snaebjornsson, J., AND Reed, D. Full-Scale Results of Wind-Induced Motion in Multi-Story Buildings. *Journal of Wind Engineering and Industrial Aerodynamics* 42, 1-3 (Oct. 1992), 1113–1123. ISSN: 01676105. DOI: 10.1016/0167-6105(92)90118-T.

[120] Laory, I., Westgate, R. J., Brownjohn, J. M. W., AND Smith, I. F. C. Temperature Variations as Loads Cases for Structural Identification. In: *Proceeding of the of 6th International Conference on Structural Health Monitoring of Intelligent Infrastructure (SHMII), Hong-Kong, 2013*. Hong Kong, 2013.

[121] Laory, I., Trinh, T. N., Smith, I. F., AND Brownjohn, J. M. Methodologies for Predicting Natural Frequency Variation of a Suspension Bridge. *Engineering Structures* 80 (Dec. 2014), 211–221. ISSN: 01410296. DOI: 10.1016/j.engstruct.2014.09.001.

[122] Yarnold, M. T., AND Moon, F. L. Temperature-Based Structural Health Monitoring Baseline for Long-Span Bridges. *Engineering Structures* 86, 0 (2015), 157 –167. ISSN: 0141-0296. DOI: 10.1016/j.engstruct.2014.12.042.

[123] Yarnold, M. T., AND Murphy, B. R. Temperature-Based Model Updating of Bridge Structures. In: American Society of Civil Engineers, 8th Feb. 2016, 909–920. ISBN: 978-0-7844-7974-2. DOI: 10.1061/9780784479742.076.

[124] Behmanesh, I., Moaveni, B., Lombaert, G., AND Papadimitriou, C. Hierarchical Bayesian Model Updating for Structural Identification. *Mechanical Systems and Signal Processing* 64-65 (Dec. 2015), 360–376. ISSN: 08883270. DOI: 10.1016/j.ymssp.2015.03.026.

[125] Ballesteros, G. C., Angelikopoulos, P., Papadimitriou, C., AND Koumoutsakos, P. Bayesian Hierarchical Models for Uncertainty Quantification in Structural Dynamics. In: American Society of Civil Engineers, 27th June 2014, 1615–1624. ISBN: 978-0-7844-1360-9. DOI: 10.1061/9780784413609.162.

[126] Behmanesh, I, AND Moaveni, B Model Updating of Structures with Temperature–Dependent Properties Using a Hierarchical Bayesian Framework. In: *SHMII-2015*. 2015.

[127] Moser, P., AND Moaveni, B. Environmental Effects on the Identified Natural Frequencies of the Dowling Hall Footbridge. *Mechanical Systems and Signal Processing* 25, 7 (Oct. 2011), 2336–2357. ISSN: 08883270. DOI: 10.1016/j.ymssp.2011.03.005.

[128] Goulet, J.-A., AND Smith, I. F. Structural Identification with Systematic Errors and Unknown Uncertainty Dependencies. *Computers & Structures* 128 (Nov. 2013), 251–258. ISSN: 00457949. DOI: 10.1016/j.compstruc.2013.07.009.

[129] Behmanesh, I., AND Moaveni, B. Bayesian FE Model Updating in the Presence of Modeling Errors. In: *Model Validation and Uncertainty Quantification, Volume 3*. Ed. by Atamturktur, H. S., Moaveni, B., Papadimitriou, C., AND Schoenherr, T. Springer International Publishing, Cham, 2014, 119–133. ISBN: 978-3-319-04551-1 978-3-319-04552-8. DOI: 10.1007/978-3-319-04552-8_12.

[130] Behmanesh, I., AND Moaveni, B. Probabilistic Identification of Simulated Damage on the Dowling Hall Footbridge through Bayesian Finite Element Model Updating. *Structural Control and Health Monitoring* 22, 3 (Mar. 2015), 463–483. ISSN: 15452255. DOI: 10.1002/stc.1684.

[131] Behmanesh, I., AND Moaveni, B. Accounting for Environmental Variability, Modeling Errors, and Parameter Estimation Uncertainties in Structural Identification. *Journal of Sound and Vibration* 374 (July 2016), 92–110. ISSN: 0022460X. DOI: 10.1016/j.jsv.2016.03.022.

[132] Behmanesh, I., Moaveni, B., AND Papadimitriou, C. Probabilistic Damage Identification of a Designed 9-Story Building Using Modal Data in the Presence of Modeling Errors. *Engineering Structures* 131 (Jan. 2017), 542–552. ISSN: 01410296. DOI: 10.1016/j.engstruct.2016.10.033.

[133] Zhang, E., Feissel, P., AND Antoni, J. A Comprehensive Bayesian Approach for Model Updating and Quantification of Modeling Errors. *Probabilistic Engineering Mechanics* 26, 4 (Oct. 2011), 550–560. ISSN: 02668920. DOI: 10.1016/j.probengmech.2011.07.001.

[134] Kennedy, M. C., AND O'Hagan, A. Bayesian Calibration of Computer Models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 63, 3 (2001), 425–464. ISSN: 1467-9868. DOI: 10.1111/1467-9868.00294.

[135] Kennedy, M. C., AND O'Hagan, A. *Supplementary Details on Bayesian Calibration of Computer Models*. University of Sheffield, 2001.

[136] Ling, Y., Mullins, J. G., AND Mahadevan, S. Options for the Inclusion of Model Discrepancy in Bayesian Calibration. In: American Institute of Aeronautics and Astronautics, 13th Jan. 2014. ISBN: 978-1-62410-312-4. DOI: 10.2514/6.2014-0122.

[137] Higdon, D., Gattiker, J., Williams, B., AND Rightley, M. Computer Model Calibration Using High-Dimensional Output. *Journal of the American Statistical Association* 103, 482 (June 2008), 570–583. ISSN: 0162-1459, 1537-274X. DOI: 10.1198/016214507000000888.

[138] Higdon, D., Kennedy, M., Cavendish, J. C., Cafeo, J. A., AND Ryne, R. D. Combining Field Data and Computer Simulations for Calibration and Prediction. *SIAM Journal on Scientific Computing* 26, 2 (Jan. 2004), 448–466. ISSN: 1064-8275, 1095-7197. DOI: 10.1137/S1064827503426693.

[139] Liu, F., Bayarri, M. J., AND Berger, J. O. Modularization in Bayesian Analysis, with Emphasis on Analysis of Computer Models. *Bayesian Analysis* 4, 1 (Mar. 2009), 119–150. ISSN: 1936-0975. DOI: 10.1214/09-BA404.

[140] Ranjan, P., Lu, W., Bingham, D., Reese, S., Williams, B. J., Chou, C.-C., Doss, F., Grosskopf, M., AND Holloway, J. P. Follow-Up Experimental Designs for Computer Models and Physical Processes. *Journal of Statistical Theory and Practice* 5, 1 (Mar. 2011), 119–136. ISSN: 1559-8608, 1559-8616. DOI: 10.1080/15598608.2011.10412055.

[141] Pearson, K. LIII. *On Lines and Planes of Closest Fit to Systems of Points in Space. Philosophical Magazine Series 6* 2, 11 (Nov. 1901), 559–572. ISSN: 1941-5982, 1941-5990. DOI: 10.1080/14786440109462720.

[142] Hastings, W. K. Monte Carlo Sampling Methods Using Markov Chains and Their Applications. *Biometrika* 57, 1 (1970), 97–109. ISSN: 0006-3444, 1464-3510. DOI: 10.1093/biomet/57.1.97.

[143] Yildirim, I. Bayesian Inference: Metropolis-Hastings Sampling. *Dept. of Brain and Cognitive Sciences, Univ. of Rochester, Rochester, NY* (2012).

[144] Besag, J., Green, P., Higdon, D., AND Mengersen, K. Bayesian Computation and Stochastic Systems. *Statistical Science* 10, 1 (Feb. 1995), 3–41. ISSN: 0883-4237. DOI: 10.1214/ss/1177010123.

[145] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., AND Teller, E. Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics* 21, 6 (June 1953), 1087–1092. ISSN: 0021-9606, 1089-7690. DOI: 10.1063/1.1699114.

[146] Arendt, P. D., Apley, D. W., Chen, W., Lamb, D., AND Gorsich, D. Improving Identifiability in Model Calibration Using Multiple Responses. *Journal of Mechanical Design* 134, 10 (2012), 100909. ISSN: 10500472. DOI: 10.1115/1.4007573.

[147] Arendt, P. D., Apley, D. W., AND Chen, W. A Preposterior Analysis to Predict Identifiability in the Experimental Calibration of Computer Models. *IIE Transactions* 48, 1 (2nd Jan. 2016), 75–88. ISSN: 0740-817X, 1545-8830. DOI: 10.1080/0740817X.2015.1064554.

[148] Jiang, Z., Apley, D., AND Chen, W. Surrogate Preposterior Analyses for Predicting and Enhancing Identifiability in Model Calibration. *International Journal for Uncertainty Quantification* 5, 4 (2015), 341–359. DOI: 10.1615/Int.J.UncertaintyQuantification.2015012627.

[149] Figueiredo, E., Radu, L., Worden, K., AND Farrar, C. R. A Bayesian Approach Based on a Markov-Chain Monte Carlo Method for Damage Detection under Unknown Sources of Variability. *Engineering Structures* 80, 0 (2014), 1 –10. ISSN: 0141-0296. DOI: 10.1016/j.engstruct.2014.08.042.

[150] Figueiredo, E., Radu, L., Westgate, R., Brownjohn, J., Cross, E., Worden, K., AND Farrar, C. Applicability of a Markov-Chain Monte Carlo Method for Damage Detection on Data from the Z-24 and Tamar Suspension Bridges. In: *Proceedings of the 6th European Workshop - Structural Health Monitoring 2012, EWSHM 2012*. Vol. 1. 2012, 747–754.

[151] Worden, K., AND Cross, E. J. Bayesian Regression Trees as Surrogate Models for the Health Monitoring of Civil Infrastructure. In: *STRUCTURAL HEALTH MONITORING 2013, VOLS 1 AND 2*. Ed. by Chang, FK. DES-TECH PUBLICATIONS, INC, 439 DUKE STREET, LANCASTER, PA 17602-4967 USA, 2013, 351–360. ISBN: 978-1-60595-115-7.

[152] Becker, W., AND Worden, K. Bayesian Sensitivity Analysis of a Bifurcating Duffing Oscillator. In: *PROCEEDINGS OF ISMA2010 - INTERNATIONAL CONFERENCE ON NOISE AND VIBRATION ENGINEERING INCLUDING USD2010*. Ed. by Sas, P and Bergen, B. Dept Mech Engn, Katholieke Univ Leuven, 2010, 5247–5259. ISBN: 978-90-73802-87-2.

[153] Gobbato, M., Kosmatka, J. B., AND Conte, J. P. A Recursive Bayesian Approach for Fatigue Damage Prognosis: An Experimental Validation at the Reliability Component Level. *MECHANICAL SYSTEMS AND SIGNAL PROCESSING* 45, 2 (4th Apr. 2014), 448–467. ISSN: 0888-3270. DOI: 10.1016/j.ymssp.2013.10.014.

[154] Enright, M., AND Frangopol, D. Condition Prediction of Deteriorating Concrete Bridges Using Bayesian Updating. *Journal of Structural Engineering* 125, 10 (1999), 1118–1125. DOI: 10.1061/(ASCE)0733-9445(1999)125:10(1118).

[155] Arroyo, D., AND Ordaz, M. Multivariate Bayesian Regression Analysis Applied to Ground-Motion Prediction Equations, Part 1: Theory and Synthetic Example. *Bulletin of the Seismological Society of America* 100, 4 (1st Aug. 2010), 1551–1567. ISSN: 0037-1106. DOI: 10.1785/0120080354.

[156] Erdogan, Y. S., Gul, M., Catbas, F. N., AND Bakir, P. G. Investigation of Uncertainty Changes in Model Outputs for Finite-Element Model Updating Using Structural Health Monitoring Data. *Journal of Structural Engineering* 140, 11 (Nov. 2014), 04014078. ISSN: 0733-9445, 1943-541X. DOI: 10.1061/(ASCE)ST.1943-541X.0001002.

[157] Erdogan, Y. S., Catbas, F. N., AND Bakir, P. G. Structural Identification (St-Id) Using Finite Element Models for Optimum Sensor Configuration and Uncertainty Quantification. *Finite Elements in Analysis and Design* 81, 0 (2014), 1 –13. ISSN: 0168-874X. DOI: 10.1016/j.finel.2013.10.009.

[158] Chen, B., AND Zang, C. A Hybrid Immune Model for Unsupervised Structural Damage Pattern Recognition. *Expert Systems with Applications* 38, 3 (Mar. 2011), 1650–1658. ISSN: 09574174. DOI: 10.1016/j.eswa.2010.07.087.

[159] Meyyappan, L, Jose, M, Dagli, C, Silva, P, AND Pottinger, H Fuzzy-Neuro System for Bridge Health Monitoring. In: *NAFIPS'2003: 22ND INTERNATIONAL CONFERENCE OF THE NORTH AMERICAN FUZZY INFORMATION PROCESSING SOCIETY - NAFIPS PROCEEDINGS*. Ed. by Walker, EL. IEEE Syst, Man & Cybernet Soc; IEEE; N Amer Fuzzy Informat Proc Soc, 2003, 8–13. ISBN: 0-7803-7918-7.

[160] Carden, E. P., AND Brownjohn, J. M. W. Fuzzy Clustering of Stability Diagrams for Vibration-Based Structural Health Monitoring. *Computer-Aided Civil and Infrastructure Engineering* 23, 5 (July 2008), 360–372. ISSN: 1093-9687, 1467-8667. DOI: 10.1111/j.1467-8667.2008.00543.x.

[161] *Kalman Filter*. 28th July 2017. URL: https://en.wikipedia.org/wiki/Kalman_filter.

[162] Ebrahimian, H., Astroza, R., AND Conte, J. P. *Parametric Identification of Hysteretic Material Constitutive Laws in Nonlinear Finite Element Models Using Extended Kalman Filter*. Department of Structural Engineering, University of California, San Diego, 2014. DOI: 10.13140/RG.2.2.12141.59364.

[163] Ebrahimian, H., Astroza, R., AND Conte, J. P. Extended Kalman Filter for Material Parameter Estimation in Nonlinear Structural Finite Element Models Using Direct Differentiation Method. *Earthquake Engineering & Structural*

Dynamics 44, 10 (Aug. 2015), 1495–1522. ISSN: 00988847. DOI: 10.1002/eqe.2532.

[164] Jin, C., Jang, S., Sun, X., Jiang, Z., AND Christenson, R. Extended Kalman Filter Based Structural Damage Detection for MR Damper Controlled Structures. In: ed. by Lynch, J. P. 20th Apr. 2016, 98031T. DOI: 10.1117/12.2218934.

[165] Jin, C., Jang, S., Sun, X., Li, J., AND Christenson, R. Damage Detection of a Highway Bridge under Severe Temperature Changes Using Extended Kalman Filter Trained Neural Network. *Journal of Civil Structural Health Monitoring* 6, 3 (July 2016), 545–560. ISSN: 2190-5452, 2190-5479. DOI: 10.1007/s13349-016-0173-8.

[166] Yang, J. N., Lin, S., Huang, H., AND Zhou, L. An Adaptive Extended Kalman Filter for Structural Damage Identification. *Structural Control and Health Monitoring* 13, 4 (July 2006), 849–867. ISSN: 1545-2255, 1545-2263. DOI: 10.1002/stc.84.

[167] Ding, Y., AND Guo, L. Structural Identification Based on Incomplete Measurements with Iterative Kalman Filter. *Structural Engineering and Mechanics* 59, 6 (25th Sept. 2016), 1037–1054. ISSN: 1225-4568. DOI: 10.12989/sem.2016.59.6.1037.

[168] Capellari, G., Eftekhar Azam, S., AND Mariani, S. Damage Detection in Flexible Plates through Reduced-Order Modeling and Hybrid Particle-Kalman Filtering. *Sensors* 16, 1 (22nd Dec. 2015), 2. ISSN: 1424-8220. DOI: 10.3390/s16010002.

[169] Ghanem, R., AND Ferro, G. Health Monitoring for Strongly Non-Linear Systems Using the Ensemble Kalman Filter. *Structural Control and Health Monitoring* 13, 1 (Jan. 2006), 245–259. ISSN: 1545-2255, 1545-2263. DOI: 10.1002/stc.139.

[170] Lei, Y., Chen, F., AND Zhou, H. An Algorithm Based on Two-Step Kalman Filter for Intelligent Structural Damage Detection. *Structural Control and Health Monitoring* 22, 4 (Apr. 2015), 694–706. ISSN: 15452255. DOI: 10.1002/stc.1712.

[171] Goulet, J. A., AND Smith, I. F. C. Predicting the Usefulness of Monitoring for Identifying the Behavior of Structures. *Journal of Structural Engineering* 139, 10 (Oct. 2013), 1716–1727. ISSN: 0733-9445, 1943-541X. DOI: 10.1061/(ASCE)ST.1943-541X.0000577.

[172] Goulet, J.-A., Kripakaran, P., AND Smith, I. F. C. Multimodel Structural Performance Monitoring. *Journal of Structural Engineering* 136, 10 (Oct. 2010), 1309–1318. ISSN: 0733-9445, 1943-541X. DOI: 10.1061/(ASCE)ST.1943-541X.0000232.

[173] Popper, K. R. *The Logic of Scientific Discovery*. Repr. 2008 (twice). Routledge classics. Routledge, London, 2008. 513 pp. ISBN: 978-0-415-27844-7.

[174] Goulet, J. A., AND Smith, I. F. C. Performance-Driven Measurement System Design for Structural Identification. *Journal of Computing in Civil Engineering* 27, 4 (July 2013), 427–436. ISSN: 0887-3801, 1943-5487. DOI: 10.1061/(ASCE)CP.1943-5487.0000250.

[175] Smith, I. F. C., Goulet, J.-A., AND Laory, I. Structural Identification Methods for Full-Scale Bridges. In: American Society of Civil Engineers, 30th Apr. 2013, 319–327. ISBN: 978-0-7844-1284-8. DOI: 10.1061/9780784412848.029.

[176] Goulet, J. Probabilistic Model Falsification for Infrastructure Diagnosis. PhD. Lausanne EPFL, 2012.

[177] Fang, S. E., AND Perera, R A Response Surface Methodology Based Damage Identification Technique. *Smart Materials and Structures* 18, 6 (June 2009), 065009. ISSN: 0964-1726, 1361-665X. DOI: 10.1088/0964-1726/18/6/065009.

[178] Fang, X., Luo, H., AND Tang, J. Structural Damage Detection Using Neural Network with Learning Rate Improvement. *Comput. Struct.* 83, 25-26 (Sept. 2005), 2150–2161. ISSN: 0045-7949. DOI: 10.1016/j.compstruc.2005.02.029.

[179] Myers, R. H., Montgomery, D. C., AND Anderson-Cook, C. M. *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*. Fourth edition. Wiley series in probability and statistics. Wiley, Hoboken, New Jersey, 2016. 825 pp. ISBN: 978-1-118-91601-8.

[180] Farrar, C., Baker, W., Bell, T., Cone, K., Darling, T., Duffey, T., Eklund, A., AND Migliori, A. *Dynamic Characterization and Damage Detection in the I-40 Bridge over the Rio Grande*. LA–12767-MS, 10158042. 1st June 1994. DOI: 10.2172/10158042.

[181] Huang, Q., Xu, Y., AND Liu, H. An Efficient Algorithm for Simultaneous Identification of Time-Varying Structural Parameters and Unknown Excitations of a Building Structure. *Engineering Structures* 98 (Sept. 2015), 29–37. ISSN: 01410296. DOI: 10.1016/j.engstruct.2015.04.019.

[182] Li, P. J., Xu, D. W., AND Zhang, J. Probability-Based Structural Health Monitoring Through Markov Chain Monte Carlo Sampling. *International Journal of Structural Stability and Dynamics* (17th Aug. 2015), 1550039. ISSN: 0219-4554, 1793-6764. DOI: 10.1142/S021945541550039X.

[183] Rasmussen, C. E., AND Williams, C. K. I. *Gaussian Processes for Machine Learning*. Adaptive computation and machine learning. MIT Press, Cambridge, Mass, 2006. ISBN: 978-0-262-18253-9.

[184] Lophaven, S. N., Nielsen, H. B., AND Søndergaard, J. *DACE - A MATLAB Kriging Toolbox*. 2002.

[185] Conti, S., Gosling, J. P., Oakley, J. E., AND O'Hagan, A. Gaussian Process Emulation of Dynamic Computer Codes. *Biometrika* 96, 3 (Sept. 2009), 663–676. ISSN: 0006-3444, 1464-3510. DOI: 10.1093/biomet/asp028.

[186] O'Hagan, A., AND Kingman, J. F. C. Curve Fitting and Optimal Design for Prediction. *Journal of the Royal Statistical Society. Series B (Methodological)* 40, 1 (1978), 1–42. ISSN: 00359246. JSTOR: 2984861.

[187] O'Hagan, A. Some Bayesian Numerical Analysis. *Bayesian statistics* 4, 345-363 (1992), 4–2.

[188] O'Hagan, A., Bernardo, J. M., Berger, J. O., Dawid, A. P., (eds, A. F. M. S., Kennedy, M. C., AND Oakley, J. E. Uncertainty Analysis and Other Inference Tools for Complex Computer Codes (1998).

[189] Carvalho, C. M., AND West, M. Dynamic Matrix-Variate Graphical Models. *Bayesian Analysis* 2, 1 (Mar. 2007), 69–97. ISSN: 1936-0975. DOI: 10.1214/07-BA204.

[190] McFarland, J., AND Swiler, L. Calibration and Uncertainty Analysis for Computer Simulations with Multivariate Output. In: *Large-Scale Inverse Problems and Quantification of Uncertainty*. Ed. by Biegler, L., Biros, G., Ghattas, O., Heinkenschloss, M., Keyes, D., Mallick, B., Marzouk, Y., Tenorio, L., van Bloemen Waanders, B., AND Willcox, K. John Wiley & Sons, Ltd, Chichester, UK, 12th Oct. 2010, 179–194. ISBN: 978-0-470-68585-3 978-0-470-69743-6.

[191] Sacks, J., Welch, W. J., Mitchell, T. J., AND Wynn, H. P. Design and Analysis of Computer Experiments. *Statistical Science* 4, 4 (Nov. 1989), 409–423. ISSN: 0883-4237. DOI: 10.1214/ss/1177012413.

[192] Kennedy, M. C., Anderson, C. W., Conti, S., AND O'Hagan, A. Case Studies in Gaussian Process Modelling of Computer Codes. *Reliability Engineering & System Safety* 91, 10-11 (Oct. 2006), 1301–1309. ISSN: 09518320. DOI: 10.1016/j.ress.2005.11.028.

[193] Genton, M. G. Separable Approximations of Space-Time Covariance Matrices. *Environmetrics* 18, 7 (Nov. 2007), 681–695. ISSN: 11804009, 1099095X. DOI: 10.1002/env.854.

[194] Carlin, B. P., AND Louis, T. A. *Bayes and Empirical Bayes Methods for Data Analysis*. 2nd ed. Chapman & Hall/CRC texts in statistical science series. Chapman & Hall/CRC, Boca Raton, 2000. ISBN: 978-1-58488-170-4.

[195] Gentle, J. E. *Matrix Algebra: Theory, Computations, and Applications in Statistics*. Springer texts in statistics. Springer, New York, N.Y. ; [London], 2007. 528 pp. ISBN: 978-0-387-70872-0 0-387-70872-3.

[196] Schäcke, K. *On the Kronecker Product*. 2004.

[197] Weiss, N. A., AND Holmes, P. T. *A Course in Probability*. International ed. Pearson/Addison-Wesley, Boston, Mass., 2006. 742 pp. ISBN: 978-0-321-18954-7.

[198] Choi, S.-C. T., Ding, Y., Hickernell, F. J., Jiang, L., Rugama, L. A. J., Tong, X., Zhang, Y., AND Zhou, X. GAIL - Guaranteed Automatic Integration Library in MATLAB: Documentation for Version 2.1. *Computer Science* abs/1503.06544 (2015).

[199] Haario, H., Saksman, E., AND Tamminen, J. An Adaptive Metropolis Algorithm. *Bernoulli* 7, 2 (Apr. 2001), 223–242.

[200] McKay, M. D., Beckman, R. J., AND Conover, W. J. A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics* 21, 2 (May 1979), 239. ISSN: 00401706. DOI: 10.2307/1268522. JSTOR: 1268522?origin=crossref.

[201] Dubbs, N. C., AND Moon, F. L. Comparison and Implementation of Multiple Model Structural Identification Methods. *Journal of Structural Engineering* 141, 11 (Nov. 2015), 04015042. ISSN: 0733-9445, 1943-541X. DOI: 10.1061/(ASCE)ST.1943-541X.0001284.

[202] Feller, W. *An Introduction to Probability Theory and Its Applications*. 2d ed. A Wiley publication in mathematical statistics. Wiley, New York, 1957. 2 pp. ISBN: 978-0-471-25709-7.

[203] Swanson Analysis Systems IP, I. *ANSYS, Inc. Documentation for Release 15.0*. Nov. 2013.

[204] Westgate, R. J., AND Brownjohn, J. M. W. Development of a Tamar Bridge Finite Element Model. In: Proulx, T. *Dynamics of Bridges, Volume 5*. Springer New York, New York, NY, 2011, 13–20. ISBN: 978-1-4419-9824-8 978-1-4419-9825-5.

[205] Westgate, R., Koo, K.-Y., AND Brownjohn, J. Effect of Solar Radiation on Suspension Bridge Performance. *Journal of Bridge Engineering* 20, 5 (May 2015), 04014077. ISSN: 1084-0702, 1943-5592. DOI: 10.1061/(ASCE)BE.1943-5592.0000668.

[206] *Tamar Bridge Wikipedia*. 1st Aug. 2017. URL: https://en.wikipedia.org/wiki/Tamar_Bridge.

[207] *Tamar Bridge Photo - Blog*. 1st Aug. 2017. URL: http://chalkygilbert.blogspot.co.uk/2012/07/meandering-over-tamar-going-abroad-to.html.

[208] Cross, E., Koo, K., Brownjohn, J., AND Worden, K. Long-Term Monitoring and Data Analysis of the Tamar Bridge. *Mechanical Systems and Signal Processing* 35, 1-2 (Feb. 2013), 16–34. ISSN: 08883270. DOI: 10.1016/j.ymssp.2012.08.026.

[209] Peeters, B., AND De Roeck, G. Reference-Based Stochastic Subspace Identification for Output-Only Modal Analysis. *Mechanical Systems and Signal Processing* 13, 6 (Nov. 1999), 855–878. ISSN: 08883270. DOI: 10.1006/mssp.1999.1249.

[210] Zhu, Y.-C., AND Au, S.-K. Bayesian Operational Modal Analysis with Asynchronous Data, Part I: Most Probable Value. *Mechanical Systems and Signal Processing* 98 (Jan. 2018), 652–666. ISSN: 08883270. DOI: 10.1016/j.ymssp.2017.05.027.

[211] Koo, K. Y., Brownjohn, J. M. W., List, D. I., AND Cole, R. Structural Health Monitoring of the Tamar Suspension Bridge. *Structural Control and Health Monitoring* 20, 4 (Apr. 2013), 609–625. ISSN: 15452255. DOI: 10.1002/stc.1481.

[212] De Battista, N., Brownjohn, J. M., Tan, H. P., AND Koo, K.-Y. Measuring and Modelling the Thermal Performance of the Tamar Suspension Bridge Using a Wireless Sensor Network. *Structure and Infrastructure Engineering* 11, 2

(Feb. 2015), 176–193. ISSN: 1573-2479, 1744-8980. DOI: 10.1080/15732479.2013.862727.

[213] Au, S.-K. OPERATIONAL MODAL ANALYSIS: HOW MUCH DATA DO WE NEED? In: *Proceeding of the of 6th International Conference on Structural Health Monitoring of Intelligent Infrastructure (SHMII), Hong-Kong, 2013*. 2013.

[214] Qureshi, Z. H., Ng, T. S., AND Goodwin, G. C. Optimum Experimental Design for Identification of Distributed Parameter Systems. *International Journal of Control* 31, 1 (Jan. 1980), 21–29. ISSN: 0020-7179, 1366-5820. DOI: 10.1080/00207178008961025.

[215] Shannon, C. E. A Mathematical Theory of Communication. *Bell System Technical Journal* 27, 3 (July 1948), 379–423. ISSN: 00058580. DOI: 10.1002/j.1538-7305.1948.tb01338.x.

[216] Scharf, L. L., AND Demeure, C. *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*. Addison-Wesley series in electrical and computer engineering. Digital signal processing. Addison-Wesley Pub. Co, Reading, Mass, 1991. 524 pp. ISBN: 978-0-201-19038-0.

[217] Ditlevsen, O. Stochastic Model for Joint Wave and Wind Loads on Offshore Structures. *Structural Safety* 24, 2-4 (Apr. 2002), 139–163. ISSN: 01674730. DOI: 10.1016/S0167-4730(02)00022-X.

[218] Damianou, A., AND Lawrence, N. Deep Gaussian Processes. In: *Proceedings of the Sixteenth International Workshop on Artificial Intelligence and Statistics (AISTATS)*. Ed. by Carvalho, C., AND Ravikumar, P. JMLR W&CP 31, Arizona, USA, 2013, 207–215.

[219] Jin, M., Damianou, A., Abbeel, P., AND Spanos, C. Inverse Reinforcement Learning via Deep Gaussian Process. *33rd Conference on Uncertainty in Artificial Intelligence (UAI)* (2017).

[220] Schwaighofer, A., Tresp, V., AND Yu, K. Learning Gaussian Process Kernels via Hierarchical Bayes. In: *Advances in Neural Information Processing Systems 17*. Ed. by Saul, L. K., Weiss, Y., AND Bottou, L. MIT Press, 2005, 1209–1216.

[221] Golub, G. H., AND Van Loan, C. F. *Matrix Computations*. 3rd ed. Johns Hopkins studies in the mathematical sciences. Johns Hopkins University Press, Baltimore, 1996. ISBN: 978-0-8018-5413-2 978-0-8018-5414-9.